



(19)  
Bundesrepublik Deutschland  
Deutsches Patent- und Markenamt

(10) DE 698 17 460 T2 2004.06.09

(12)

## Übersetzung der europäischen Patentschrift

(97) EP 0 884 912 B1  
(21) Deutsches Aktenzeichen: 698 17 460.7  
(96) Europäisches Aktenzeichen: 98 110 471.4  
(96) Europäischer Anmeldetag: 08.06.1998  
(97) Erstveröffentlichung durch das EPA: 16.12.1998  
(97) Veröffentlichungstag  
der Patenterteilung beim EPA: 27.08.2003  
(47) Veröffentlichungstag im Patentblatt: 09.06.2004

(51) Int Cl.: H04N 7/36  
H04N 7/26; H04N 7/46

(30) Unionspriorität:  
15065697 09.06.1997 JP

(84) Benannte Vertragsstaaten:  
DE, ES, FI, FR, GB, IT, NL, SE

(73) Patentinhaber:  
Hitachi, Ltd., Tokio/Tokyo, JP

(72) Erfinder:  
Nakaya, Yuichiro, Tokyo, JP

(74) Vertreter:  
Strehl, Schübel-Hopf & Partner, 80538 München

(54) Bezeichnung: Bildsequenzdekodierungsverfahren

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99 (1) Europäisches Patentübereinkommen).

Die Übersetzung ist gemäß Artikel II § 3 Abs. 1 IntPatÜG 1991 vom Patentinhaber eingereicht worden. Sie wurde vom Deutschen Patent- und Markenamt inhaltlich nicht geprüft.

DE 698 17 460 T2 2004.06.09

## Beschreibung

## HINTERGRUND DER ERFINDUNG

## Gebiet der Erfindung

[0001] Die vorliegende Erfindung betrifft ein Bildfolgedecodierverfahren, welches eine Zwischenrahmenvorhersage ausführt.

## Beschreibung des Stands der Technik

[0002] Bei der hochwirksamen Codierung von Bildfolgen ist die Zwischenrahmenvorhersage (Bewegungskompensation) unter Verwendung der Ähnlichkeit angrenzender Frames (im folgenden als "Rahmen" übersetzt) über die Zeit als eine hochwirksame Technik zur Datenkompression bekannt. Das heute am häufigsten verwendete Bewegungskompensationsverfahren ist die Blockanpassung mit einer Halbpxelgenauigkeit, die in den internationalen Standards H.263, MPEG1 und MPEG2 verwendet wird. Bei diesem Verfahren wird das zu codierende Bild in Blöcke segmentiert, und die horizontalen und vertikalen Komponenten der Bewegungsvektoren dieser Blöcke werden als ganzzahlige Vielfache der Hälfte des Abstands zwischen benachbarten Pixeln geschätzt. Dieser Prozeß wird unter Verwendung der folgenden Gleichung beschrieben:

[Gleichung 1]

[0003]

$$P(x, y) = R(x + u_i, y + v_i)(x, y) \in B_i, 0 \leq i < N$$

(1)

[0004] Hierbei bezeichnen  $P(x, y)$  und  $R(x, y)$  die Probenwerte (Luminanz oder Chrominanzintensität) von Pixeln, die sich an Koordinaten  $(x, y)$  im vorausgesagten Bild  $P$  des aktuellen Rahmens bzw. im Referenzbild  $R$  (das decodierte Bild eines Rahmens, der vor dem aktuellen Rahmen codiert wurde) befinden.  $x$  und  $y$  sind ganze Zahlen, und es wird angenommen, daß sich alle Pixel an Punkten befinden, an denen die Koordinatenwerte ganze Zahlen sind. Zusätzlich wird angenommen, daß die Abtastwerte der Pixel zu nicht negativen ganzen Zahlen quantisiert sind.  $N$ ,  $B_i$  und  $(u_i, v_i)$  bezeichnen die Anzahl der Blöcke im Bild, den im  $i$ -ten Block des Bilds enthaltenen Pixelsatz bzw. die Bewegungsvektoren des  $i$ -ten Blocks.

[0005] Wenn die Werte von  $u_i$  und  $v_i$  keine ganzen Zahlen sind, ist es erforderlich, den Intensitätswert an dem Punkt zu finden, an dem tatsächlich keine Pixel im Referenzbild existieren. Gegenwärtig ist die bilineare Interpolation unter Verwendung der benachbarten vier Pixel das am häufigsten verwendete Verfahren für diesen Prozeß. Dieses Interpolationsverfahren wird unter Verwendung der folgenden Gleichung beschrieben:

[Gleichung 2]

$$R\left(x + \frac{p}{d}, y + \frac{q}{d}\right) = ((d - q)((d - p)R(x, y) + pR(x + 1, y)) + q((d - p)R(x, y + 1) + pR(x + 1, y + 1))) / d^2 \quad (2)$$

[0006] Hierbei ist  $d$  eine positive ganze Zahl und sind  $p$  und  $q$  kleiner als  $d$ , jedoch nicht kleiner als 0. "/" bezeichnet eine ganzzahlige Division, bei der das Ergebnis einer normalen Division (Division unter Verwendung realer Zahlen) gegen die nächste ganze Zahl gerundet wird.

[0007] Ein Beispiel des Aufbaus eines H.263-Videocodierers ist in Fig. 1 dargestellt. Als Codieralgorithmus verwendet H.263 ein Hybridcodierverfahren (adaptives Zwischenrahmen/Innenrahmen-Codierverfahren), das eine Kombination der Blockanpassung und der DCT (diskrete Kosinustransformation) ist. Ein Subtrahierer 102 berechnet die Differenz zwischen dem Eingangsbild (aktuellen Rahmenbasisbild) 101 und dem Ausgangsbild 113 (später angesprochen) des Zwischenrahmen/Innenrahmen-Codierwählers 119 und gibt dann ein Fehlerbild 103 aus. Dieses Fehlerbild wird in einem Quantisierer 105 quantisiert, nachdem es in einem DCT-Wandler 104 in DCT-Koeffizienten umgewandelt wurde, und bildet dann quantisierte DCT-Koeffizienten 106. Diese quantisierten DCT-Koeffizienten werden über den Kommunikationskanal übertragen, während sie gleichzeitig zum Synthetisieren des vorausgesagten Zwischenrahmenbilds im Codierer verwendet werden. Die Prozedur zum Synthetisieren des vorausgesagten Bilds wird als nächstes erklärt. Die vorstehend erwähnten quantisierten DCT-Koeffizienten 106 bilden das rekonstruierte Fehlerbild 110 (das dem rekonstruierten Fehlerbild auf der Empfangsseite gleicht), nachdem sie einen Dequantisierer 108 und einen inversen DCT-Wandler 109 durch-

DE 698 17 460 T2 2004.06.09

laufen haben. Dieses rekonstruierte Fehlerbild und das Ausgangsbild 113 des Zwischenrahmen/Innenrahmen-Codierwählers 119 werden am Addierer 111 addiert, und das decodierte Bild 112 des aktuellen Rahmens (das dem decodierten Bild des auf der Empfängerseite rekonstruierten aktuellen Rahmens gleicht) wird erhalten. Dieses Bild wird in einem Rahmenspeicher 114 gespeichert und um eine Zeit verzögert, die dem Rahmenintervall gleicht. Dementsprechend gibt der Rahmenspeicher 114 am aktuellen Punkt das decodierte Bild 115 des vorausgegangenen Rahmens aus. Dieses decodierte Bild des vorausgegangenen Rahmens und das Originalbild 101 des aktuellen Rahmens werden in den Blockanpassungsabschnitt 116 eingegeben, und es wird zwischen diesen Bildern eine Blockanpassung ausgeführt. Beim Blockanpassungsprozeß wird das Originalbild des aktuellen Rahmens in mehrere Blöcke segmentiert, und das vorausgesagte Bild 117 des aktuellen Rahmens wird durch Extrahieren des diesen Blöcken am meisten ähnelnden Abschnitts aus dem decodierten Bild des vorausgegangenen Rahmens synthetisiert. Es ist bei diesem Prozeß erforderlich, die Bewegung zwischen dem vorhergehenden Rahmen und dem aktuellen Rahmen für jeden Block zu schätzen. Der im Bewegungsschätzprozeß geschätzte Bewegungsvektor für jeden Block wird als Bewegungsvektordaten 120 zur Empfängerseite übertragen. Auf der Empfängerseite wird das gleiche Voraussagebild wie auf der Senderseite unter Verwendung der Bewegungsvektorinformationen und des Decodierbilds des vorausgegangenen Rahmens synthetisiert. Das Voraussagebild 117 wird zusammen mit einem "0"-Signal 118 in den Zwischenrahmen/Innenrahmen-Codierwähler 119 eingegeben. Dieser Schalter 119 wählt die Zwischenrahmencodierung oder die Innenrahmencodierung durch Wählen von einem dieser Eingaben aus. Die Zwischenrahmencodierung wird ausgeführt, wenn das Voraussagebild 117 gewählt wird (dieser Fall ist in Fig. 2 dargestellt). Wenn andererseits das "0"-Signal gewählt wird, wird die Innenrahmencodierung ausgeführt, weil das Eingangsbild selbst in DCT-Koeffizienten umgewandelt und an den Kommunikationskanal ausgegeben wird. Damit die Empfängerseite das codierte Bild richtig rekonstruiert, muß der Empfänger informiert werden, ob auf der Senderseite eine Innenrahmencodierung oder eine Zwischenrahmencodierung ausgeführt wurde. Folglich wird ein Identifizierkennzeichen 121 an die Kommunikationsschaltung ausgegeben. Schließlich wird ein H.263-codierter Bitstrom 123 durch Multiplexen der quantisierten DCT-Koeffizienten, Bewegungsvektoren und der Zwischenrahmen/Innenrahmen-Identifizierkennzeicheninformationen in einem Multiplexer 122 gewonnen.

[0008] Der Aufbau eines Decodierers 200 zum Empfangen des vom Codierer aus Fig. 1 ausgegebenen codierten Bitstroms ist in Fig. 2 dargestellt. Der H.263-codierte Bitstrom 217, der empfangen wird, wird im Demultiplexer 216 in quantisierte DCT-Koeffizienten 201, in Bewegungsvektordaten 202 und ein Zwischenrahmen/Innenrahmen-Identifizierkennzeichen 203 demultiplext. Die quantisierten DCT-Koeffizienten 201 werden zu einem decodierten Fehlerbild 206, nachdem sie durch einen inversen Quantisierer 204 und einen inversen DCT-Wandler 205 verarbeitet wurden. Dieses decodierte Fehlerbild wird in einem Addierer 207 zum Ausgangsbild 215 des Zwischenrahmen/Innenrahmen-Codierwählers 214 addiert, und die Summe dieser Bilder wird als das decodierte Bild 208 ausgegeben. Der Ausgang des Zwischenrahmen/Innenrahmen-Codierwählers wird entsprechend dem Zwischenrahmen/Innenrahmen-Identifizierkennzeichen 203 umgeschaltet. Ein Voraussagebild 212, das verwendet wird, wenn eine Zwischenrahmencodierung ausgeführt wird, wird im Voraussagebildsynthetisierer 211 synthetisiert. In diesem Synthetisierer wird die Position der Blöcke im decodierten Bild 210 des im Rahmenspeicher 209 gespeicherten vorhergehenden Rahmens entsprechend den Bewegungsvektordaten 202 verschoben. Andererseits gibt der Zwischenrahmen/Innenrahmen-Codierwähler das "0"-Signal 213 unverändert zur Innenrahmencodierung aus.

[0009] EP-A-0 712 249 beschreibt Pixelinterpolation unter Verwendung von positivem Runden weg von Null.

#### ZUSAMMENFASSUNG DER ERFINDUNG

[0010] Das durch H.263 codierte Bild besteht aus einer Luminanzinformationen enthaltenden Luminanzebene (Y-Ebene) und zwei Chrominanzinformationen enthaltenden Chrominanzebenen (U-Ebene und V-Ebene). Es ist dabei charakteristisch, daß die Y-Ebene  $2m$  Pixel horizontal und  $2n$  Pixel vertikal aufweist und die U- und die V-Ebene  $m$  Pixel horizontal und  $n$  Pixel vertikal aufweisen, wenn das Bild  $2m$  Pixel in horizontaler Richtung und  $2n$  Pixel in vertikaler Richtung ( $m$  und  $n$  sind positive ganze Zahlen) aufweist. Die niedrige Auflösung in der Chrominanzebene ist darauf zurückzuführen, daß das menschliche Sehsystem eine verhältnismäßig geringe Sehfähigkeit hinsichtlich räumlicher Änderungen der Chrominanz aufweist. Wenn ein solches Bild als Eingabe vorliegt, führt H.263 eine Codierung und Decodierung in als Makroblöcke bezeichneten Blockeinheiten aus. Der Aufbau eines Makroblocs ist in Fig. 3 dargestellt. Der Makroblock besteht aus drei Blöcken, nämlich einem Y-Block, einem U-Block und einem V-Block. Die Größe des die Luminanzinformationen enthaltenden Y-Blocks 301 beträgt  $16 \times 16$  Pixel, und die Größe des die Chrominanzinformationen enthaltenden U-Blocks 302 und V-Blocks 303 beträgt  $8 \times 8$  Pixel.

[0011] Bei H.263 wird eine Blockanpassung mit Halbpixelgenauigkeit auf jeden Block angewendet. Wenn der geschätzte Bewegungsvektor demgemäß als  $(u, v)$  definiert ist, sind  $u$  und  $v$  beide ganzzahlige Vielfache der Hälfte des Abstands zwischen Pixeln. Mit anderen Worten wird  $1/2$  als die minimale Einheit verwendet. Die Konfiguration des für die Intensitätswerte verwendeten Interpolationsverfahrens (nachstehend werden die In-

DE 698 17 460 T2 2004.06.09

Intensitätswerte für die "Luminanz" und die "Chrominanz" mit dem allgemeinen Begriff "Intensitätswert" bezeichnet) ist in Fig. 4 dargestellt. Wenn die in Gleichung 2 beschriebene Interpolation ausgeführt wird, werden die Quotienten der Division gegen die nächste ganze Zahl abgerundet, und wenn der Quotient weiter einen halbzahlgigen Wert hat (also 0,5 zu einer ganzen Zahl addiert), wird ein Abrunden zur nächsten ganzen Zahl von Null weg ausgeführt. Wenn in Fig. 4 mit anderen Worten die Intensitätswerte für 401, 402, 403, 404 La, Lb, Lc bzw. Ld sind (La, Lb, Lc und Ld sind nicht negative ganze Zahlen), lassen sich die interpolierten Intensitätswerte la, lb, lc und ld (la, lb, lc und ld sind nicht negative ganze Zahlen) an den Positionen 405, 406, 407, 408 durch die folgende Gleichung ausdrücken:

[Gleichung 3]

[0012]

$$la = La$$

$$lb = [(La + Lb + 1)/2]$$

$$lc = [(La + Lc + 1)/2]$$

$$ld = [(La + Lb + Lc + Ld + 2)/4] \quad (3)$$

wobei "[ ]" das Abrunden zur nächsten ganzen Zahl gegen 0 bezeichnet (wobei der Bruchteil verworfen wird). Die Erwartung der durch dieses Runden auf ganze Zahlen bewirkten Fehler wird folgendermaßen geschätzt: Es wird angenommen, daß die Wahrscheinlichkeiten, daß der Intensitätswert an den Positionen 405, 406, 407 und 408 aus Fig. 4 verwendet wird, alle 25 Prozent sind. Wenn der Intensitätswert la für die Position 405 gefunden wird, ist der Rundungsfehler klarerweise null. Wenn weiterhin der Intensitätswert lb für die Position 406 gefunden wird, ist der Fehler null, wenn La + Lb eine gerade Zahl ist, und der Fehler ist 1/2, wenn La + Lb eine ungerade Zahl ist. Falls die Wahrscheinlichkeit, daß La + Lb eine gerade Zahl ist, und die Wahrscheinlichkeit, daß La + Lb eine ungerade Zahl ist, beide 50 Prozent sind, ist die Erwartung für den Fehler  $0 \times 1/2 + 1/2 \times 1/2 = 1/4$ . Wenn weiterhin der Intensitätswert lc für die Position 407 gefunden wird, ist die Erwartung für den Fehler ebenso wie bei lb 1/4. Wenn der Intensitätswert ld für die Position 408 gefunden wird, ist der Fehler, wenn der Rest von La + Lb + Lc + Ld geteilt durch vier 0, 1, 2 bzw. 3 ist, 0, -1/4, 1/2 bzw. 1/4. Falls angenommen wird, daß die Wahrscheinlichkeiten, daß der Rest 0, 1, 2 und 3 ist, alle gleich sind (also 25 Prozent), ist die Erwartung für den Fehler  $0 \times 1/4 - 1/4 \times 1/4 + 1/2 \times 1/4 + 1/4 \times 1/4 = 1/8$ . Wie vorstehend beschrieben wurde, ist die endgültige Erwartung für den Fehler unter der Annahme, daß die Wahrscheinlichkeiten, daß die an den Positionen 405-408 verwendeten Intensitätswerte alle gleich sind,  $0 \times 1/4 + 1/4 \times 1/4 + 1/4 \times 1/4 + 1/8 \times 1/4 = 5/32$ . Dies weist darauf hin, daß jedesmal dann, wenn eine Bewegungskompensation durch eine Blockanpassung ausgeführt wird, ein Fehler von 5/32 im Pixelintensitätswert auftritt. Allgemein kann beim Codieren mit einer niedrigen Rate keine ausreichende Anzahl von Bits für das Codieren der Zwischenrahmen-Fehlerdifferenz verwendet werden, so daß die quantisierte Schrittgröße des DCT-Koeffizienten leicht zu groß wird. Dementsprechend werden infolge der Bewegungskompensation auftretende Fehler nur dann korrigiert, wenn sie sehr groß ist. Wenn eine Zwischenrahmen-codierung kontinuierlich ausgeführt wird, ohne daß in dieser Umgebung eine Innenrahmen-codierung ausgeführt wird, neigen die Fehler zum Akkumulieren und führen zu ungünstigen Wirkungen auf das rekonstruierte Bild.

[0013] Wie vorstehend erklärt wurde, liegt in der vertikalen und der horizontalen Richtung der Chrominanzebene jeweils die Hälfte der Anzahl der Pixel. Daher wird für die Bewegungsvektoren des U-Blocks und des V-Blocks die Hälfte des Werts des Bewegungsvektors für den Y-Block für die vertikale und die horizontale Komponente verwendet. Weil die horizontale und die vertikale Komponente des Bewegungsvektors für den Bewegungsvektor des Y-Blocks ganzzahlige Vielfache von 1/2 sind, erscheinen die Bewegungsvektorkomponenten für den U-Block und den V-Block als ganzzahlige Vielfache von 1/4 (Viertelpixelgenauigkeit), falls eine gewöhnliche Division ausgeführt wird. Infolge der hohen Komplexität der Berechnung des Intensitätsinterpolationsprozesses für Bewegungsvektoren mit Viertelpixelgenauigkeit werden die Bewegungsvektoren für die U- und V-Blöcke bei H.263 jedoch zur Halbpixelgenauigkeit gerundet. Das bei H.263 verwendete Rundungsverfahren ist das folgende: Nach der vorstehend beschriebenen Definition bezeichnet (u, v) den Bewegungsvektor des Makroblocks (der dem Bewegungsvektor für den Y-Block gleicht). Unter der Annahme, daß r eine ganze Zahl ist und daß s eine nicht negative ganze Zahl kleiner als 4 ist, kann u/2 als  $u/2 = r + s/4$  umgeschrieben werden. Wenn s 0 oder 2 ist, ist kein Runden erforderlich, weil u/2 bereits ein ganzzahliges Vielfaches von 1/2 ist. Wenn s jedoch gleich 1 oder 3 ist, wird der Wert von s auf 2 gerundet. Durch Erhöhen der Wahrscheinlichkeit, daß s unter Verwendung dieses Rundungsverfahrens den Wert 2 annimmt, kann die Filterwirkung der Bewegungskompensation verstärkt werden. Wenn die Wahrscheinlichkeiten, daß der Wert von s vor dem Runden 0, 1, 2

DE 698 17 460 T2 2004.06.09

und 3 ist, alle 25 Prozent sind, sind die Wahrscheinlichkeiten, daß s nach dem Runden 0 oder 2 ist, 25 Prozent bzw. 75 Prozent. Der vorstehend erklärte Prozeß, der sich auf die horizontale Komponente u des Bewegungsvektors bezieht, läßt sich auch auf die vertikale Komponente v anwenden. Dementsprechend beträgt die Wahrscheinlichkeit für die Verwendung des Intensitätswerts an der Position 401 im U-Block und im V-Block  $1/4 \times 1/4 = 1/16$ , und die Wahrscheinlichkeiten für die Verwendung des Intensitätswerts der Positionen 402 und 403 betragen beide  $1/4 \times 3/4 = 3/16$ , während die Wahrscheinlichkeit für die Verwendung des Intensitätswerts an der Position 404  $3/4 \times 3/4 = 9/16$  ist. Unter Verwendung des gleichen Verfahrens wie vorstehend erwähnt, beträgt die Erwartung für den Fehler des Intensitätswerts  $0 \times 1/16 + 1/4 \times 3/16 + 1/4 \times 3/16 + 1/8 \times 9/16 = 21/128$ . Wie vorstehend für den Y-Block erklärt wurde, tritt das Problem akkumulierter Fehler auf, wenn die Zwischenrahmencodierung kontinuierlich ausgeführt wird.

[0014] Wie vorstehend erwähnt wurde, tritt für Bildsequenzcodier- und -decodierverfahren, bei denen eine Zwischenrahmenvoraussage ausgeführt wird und die Luminanz- oder Chrominanzintensität quantisiert wird, das Problem akkumulierter Rundungsfehler auf. Diese Rundungsfehler werden erzeugt, wenn der Luminanz- oder Chrominanzintensitätswert während der Erzeugung des Zwischenrahmen-Voraussagebilds quantisiert wird.

[0015] Angesichts der vorstehenden Probleme besteht eine Aufgabe dieser Erfindung darin, die Qualität des rekonstruierten Bilds durch Verhindern einer Fehlerakkumulation zu verbessern.

[0016] Zur Lösung dieser Aufgabe wird die Fehlerakkumulation dadurch verhindert, daß das Auftreten von Fehlern begrenzt oder eine Operation zum Aufheben der aufgetretenen Fehler durchgeführt wird.

[0017] Die Aufgabe wird durch das Verfahren nach Anspruch 1 und den Decoder nach Anspruch 2 gelöst. Die Unteransprüche betreffen bevorzugte Ausführungsformen davon.

#### KURZBESCHREIBUNG DER ZEICHNUNG

[0018] Fig. 1 ist ein Blockdiagramm, in dem das Layout des H.263-Bildcodierers dargestellt ist.

[0019] Fig. 2 ist ein Blockdiagramm, in dem das Layout des H.263-Bilddecodierers dargestellt ist.

[0020] Fig. 3 zeigt den Aufbau des Makroblocks.

[0021] Fig. 4 zeigt den Prozeß der Interpolation von Intensitätswerten für die Blockanpassung mit Halbpixelgenauigkeit.

[0022] Fig. 5 zeigt eine codierte Bildfolge.

[0023] Fig. 6 ist ein Blockdiagramm, in dem eine Software-Bildcodiervorrichtung dargestellt ist.

[0024] Fig. 7 ist ein Blockdiagramm, in dem eine Software-Bilddecodiervorrichtung dargestellt ist.

[0025] Fig. 8 ist ein Flußdiagramm, in dem ein Beispiel der Verarbeitung in der Software-Bildcodiervorrichtung dargestellt ist.

[0026] Fig. 9 ist ein Flußdiagramm, in dem ein Beispiel der Codiermodus-Entscheidungsverarbeitung für die Software-Bildcodiervorrichtung dargestellt ist.

[0027] Fig. 10 ist ein Flußdiagramm, in dem ein Beispiel der Bewegungsschätz- und der Bewegungskompensationsverarbeitung in der Software-Bildcodiervorrichtung dargestellt ist.

[0028] Fig. 11 ist ein Flußdiagramm, in dem die Verarbeitung in der Software-Bilddecodiervorrichtung dargestellt ist.

[0029] Fig. 12 ist ein Flußdiagramm, in dem ein Beispiel der Bewegungskompensationsverarbeitung in der Software-Bilddecodiervorrichtung dargestellt ist.

[0030] Fig. 13 zeigt ein Beispiel eines Speichermediums, auf dem ein codierter Bitstrom aufgezeichnet wird, der durch ein Codierverfahren erzeugt wurde, das I-, (P+)- und (P-)-Rahmen enthaltende Bitströme ausgibt.

[0031] Fig. 14 ist ein Darstellungssatz, in dem spezifische Beispiele von Vorrichtungen gezeigt sind, wobei ein Codierverfahren verwendet wird, bei dem (P+)- und (P-)-Rahmen gemeinsam auftreten.

[0032] Fig. 15 zeigt ein Beispiel eines Speichermediums, auf dem ein codierter Bitstrom aufgezeichnet wird, der durch ein Codierverfahren erzeugt wurde, das I-, B-, (P+)- und (P-)-Rahmen enthaltende Bitströme ausgibt.

[0033] Fig. 16 ist ein Blockdiagramm eines Beispiels einer Blockanpassungseinheit, die in einer Vorrichtung enthalten ist, wobei ein Codierverfahren verwendet wird, bei dem (P+)- und (P-)-Rahmen gemeinsam auftreten.

[0034] Fig. 17 ist ein Blockdiagramm des in einer Vorrichtung zum Decodieren von Bitströmen, die durch ein Codierverfahren codiert wurden, bei dem (P+)- und (P-)-Rahmen gemeinsam auftreten, enthaltenen Voraussagebildsynthetisierers.

#### DETAILLIERTE BESCHREIBUNG DER BEVORZUGTEN AUSFÜHRUNGSFORMEN

[0035] Zuerst muß erwogen werden, unter welchen Umständen die akkumulierten Rundungsfehler, die im "Stand der Technik" auftreten, berücksichtigt werden müssen. Ein Beispiel einer Bildfolge, die durch Codier-

DE 698 17 460 T2 2004.06.09

verfahren codiert wurde, welche sowohl eine unidirektionale Voraussage als auch eine bidirektionale Voraussage ausführen können, wie es bei MPEG.1, MPEG.2 und H.263 der Fall ist, ist in Fig. 5 dargestellt. Ein Bild 501 wird durch eine Innenrahmencodierung codiert und als ein I-Rahmen bezeichnet. Dagegen werden die Bilder 503, 505, 507, 509 als P-Rahmen bezeichnet und durch unidirektionale Zwischenrahmencodierung unter Verwendung des vorausgegangenen I- oder P-Rahmens als das Referenzbild codiert. Wenn dementsprechend beispielsweise das Bild 505 codiert wird, wird das Bild 503 als Referenzbild verwendet und eine Zwischenrahmenvoraussage ausgeführt. Die Bilder 502, 504, 506 und 508 werden als B-Rahmen bezeichnet, und eine bidirektionale Zwischenrahmenvoraussage wird unter Verwendung des vorausgegangenen und des folgenden I- oder P-Rahmens ausgeführt. Der B-Rahmen ist dadurch gekennzeichnet, daß er nicht als ein Referenzbild verwendet wird, wenn die Zwischenrahmenvoraussage ausgeführt wird. Weil in I-Rahmen keine Bewegungskompensation vorgenommen wird, tritt der durch die Bewegungskompensation hervorgerufene Rundungsfehler nicht auf. Dagegen wird bei den P-Rahmen nicht nur die Bewegungskompensation ausgeführt, sondern der P-Rahmen wird auch als ein Referenzbild von anderen P- oder B-Rahmen verwendet, so daß er eine Ursache für akkumulierte Rundungsfehler sein kann. Bei den B-Rahmen wird die Bewegungskompensation andererseits so ausgeführt, daß die Wirkung akkumulierter Rundungsfehler im rekonstruierten Bild auftritt. Weil B-Rahmen jedoch nicht als Referenzbilder verwendet werden, können B-Rahmen keine Quelle akkumulierter Rundungsfehler sein. Falls demgemäß akkumulierte Rundungsfehler im P-Rahmen verhindert werden können, können die ungünstigen Wirkungen von Rundungsfehlern in der Gesamtbildfolge verringert werden. Bei H.263 existiert ein Rahmen zum Codieren eines P-Rahmens und eines B-Rahmens und wird als PB-Rahmen bezeichnet (beispielsweise können die Rahmen 503 und 504 beide als ein PB-Rahmen codiert werden). Falls die beiden kombinierten Rahmen als getrennte Rahmen betrachtet werden, kann das gleiche Prinzip angewendet werden, das vorstehend erwähnt wurde. Falls mit anderen Worten Gegenmaßnahmen gegen Rundungsfehler für den P-Rahmenteil innerhalb eines PB-Rahmens ergriffen werden, kann die Akkumulation von Fehlern verhindert werden.

[0036] Rundungsfehler treten während der Interpolation von Intensitätswerten auf, wenn ein aus einer normalen Division (Division, deren Operationsergebnis eine reale Zahl ist) erhaltener Wert eine Halbzahl ist (0,5 zu einer ganzen Zahl addiert), und dieses Ergebnis wird dann zur nächsten ganzen Zahl von Null weg aufgerundet. Wenn beispielsweise durch 4 dividiert wird, um einen interpolierten Intensitätswert zu finden, haben die Rundungsfehler für die Fälle, in denen der Rest 1 und 3 ist, gleiche Absolutwerte, jedoch unterschiedliche Vorzeichen. Folglich werden die von diesen zwei Fällen hervorgerufenen Rundungsfehler aufgehoben, wenn die Erwartung für die Rundungsfehler berechnet wird (mit anderen Worten werden die in diesen Fällen hervorgerufenen Rundungsfehler beim Dividieren durch eine positive ganze Zahl d' aufgehoben, wenn der Rest t und d' - t ist). Wenn der Rest jedoch 2 ist und das Ergebnis der normalen Division mit anderen Worten eine Halbzahl ist, können die Rundungsfehler nicht aufgehoben werden, und es ergeben sich akkumulierte Fehler. Zum Lösen dieses Problems kann ein Verfahren verwendet werden, das die Verwendung von zwei Rundungsverfahren ermöglicht. Die zwei hier verwendeten Rundungsverfahren sind ein Rundungsverfahren, das Halbzahlen von Null weg rundet, und ein Rundungsverfahren, das Halbzahlen gegen null rundet. Durch Kombinieren der Verwendung dieser zwei Rundungsverfahren können die Rundungsfehler aufgehoben werden. Nachstehend wird das Rundungsverfahren, das das Ergebnis der normalen Division gegen die nächste ganze Zahl rundet und halbzahlige Werte von Null weg rundet, als "positives Runden" bezeichnet. Weiterhin wird das Rundungsverfahren, das das Ergebnis der normalen Division gegen die nächste ganze Zahl rundet und halbzahlige Werte gegen null rundet, als "negatives Runden" bezeichnet. Der Prozeß des bei der Blockanpassung mit Halbpixelgenauigkeit verwendeten positiven Rundens ist in Gleichung 3 dargestellt. Wenn stattdessen ein negatives Runden verwendet wird, kann diese Gleichung wie nachstehend dargestellt umgeschrieben werden.

[Gleichung 4]

[0037]

 $I_a = L_a$  $I_b = [(L_a + L_b)/2]$  $I_c = [(L_a + L_c)/2]$  $I_d = [(L_a + L_b + L_c + L_d + 1)/4]$  (4)

[0038] Nachstehend werden Bewegungskompensationsverfahren, die zur Synthese von Zwischenrahmenvoraussagebildern ein positives und negatives Runden ausführen, als "Bewegungskompensation unter Verwendung eines positiven Rundens" bzw. als "Bewegungskompensation unter Verwendung eines negativen

DE 698 17 460 T2 2004.06.09

Rundens" bezeichnet. Weiterhin wird bei P-Rahmen, bei denen eine Blockanpassung mit einer Halbpixelgenauigkeit zur Bewegungskompensation verwendet wird, ein Rahmen, bei dem ein positives Runden verwendet wird, als ein "(P+)-Rahmen" bezeichnet, und ein Rahmen, bei dem ein negatives Runden verwendet wird, als ein "(P-)-Rahmen" bezeichnet (nach dieser Definition sind die P-Rahmen bei H.263 alle (P+)-Rahmen). Die Erwartungen für die Rundungsfehler bei (P+)- und (P-)-Rahmen haben gleiche Absolutwerte, jedoch verschiedene Vorzeichen. Dementsprechend kann die Akkumulation von Rundungsfehlern verhindert werden, wenn (P+)-Rahmen und (P-)-Rahmen alternierend entlang der Zeitachse angeordnet sind. Falls in dem Beispiel in Fig. 5 die Rahmen 503 und 507 als (P+)-Rahmen festgelegt sind und die Rahmen 505 und 509 als (P-)-Rahmen festgelegt sind, kann dieses Verfahren implementiert werden. Das alternierende Auftreten von (P+)-Rahmen und (P-)-Rahmen führt zur Verwendung eines (P+)-Rahmens und eines (P-)-Rahmens bei der bidirektionalen Voraussage für B-Rahmen. Der Durchschnitt des Vorwärtsvoraussagebilds (also des Voraussagebilds, das unter Verwendung des Rahmens 503 synthetisiert wird, wenn der Rahmen 504 in Fig. 5 codiert wird) und des Rückwärtsvoraussagebilds (also des Voraussagebilds, das unter Verwendung des Rahmens 505 synthetisiert wird, wenn der Rahmen 504 in Fig. 5 codiert wird) wird häufig zur Synthese des Voraussagebilds für B-Rahmen verwendet. Dies bedeutet, daß die Verwendung eines (P+)-Rahmens (der einen positiven Wert für die Erwartung des Rundungsfehlers hat) und eines (P-)-Rahmens (der einen negativen Wert für die Erwartung des Rundungsfehlers hat) bei der bidirektionalen Voraussage für einen B-Rahmen beim Aufheben der Wirkungen von Rundungsfehlern wirksam ist. Wie vorstehend erwähnt wurde, ist der Rundungsprozeß im B-Rahmen keine Ursache für eine Fehlerakkumulation. Dementsprechend tritt selbst dann kein Problem auf, wenn das gleiche Rundungsverfahren auf alle B-Rahmen angewendet wird. Beispielsweise tritt selbst dann keine ernsthafte Verschlechterung der decodierten Bilder auf, wenn eine Bewegungskompensation unter Verwendung eines positiven Rundens für alle B-Rahmen 502, 504, 506 und 508 in Fig. 5 ausgeführt wird. Vorzugsweise wird nur ein Typ eines Rundens für einen Rahmen ausgeführt, um den B-Rahmen-Decodierprozeß zu vereinfachen.

[0039] Ein Blockanpassungsabschnitt 1600 eines Bildcodierers gemäß dem vorstehend beschriebenen Bewegungskompensationsverfahren unter Verwendung von Mehrfachrundungsverfahren ist in Fig. 16 dargestellt. Zahlen, die mit jenen in anderen Zeichnungsbestandteilen identisch sind, geben die gleichen Teile an. Durch Ersetzen des Blockanpassungsabschnitts 116 aus Fig. 1 durch 1600 können Mehrfachrundungsverfahren verwendet werden. Eine Bewegungsschätzverarbeitung zwischen dem Eingangs- und dem decodierten Bild des vorausgegangenen Rahmens wird in einem Bewegungsschätzer 1601 ausgeführt. Als Ergebnis werden Bewegungsinformationen 120 ausgegeben. Diese Bewegungsinformationen werden bei der Synthese des Voraussagebilds in einem Voraussagebildsynthesizer 1603 verwendet. Eine Rundungsverfahren-Bestimmungsvorrichtung 1602 bestimmt, ob ein positives Runden oder ein negatives Runden als das Rundungsverfahren für den gegenwärtig codierten Rahmen zu verwenden ist. Informationen 1604, die sich auf das bestimmte Rundungsverfahren beziehen, werden in den Voraussagebildsynthesizer 1603 eingegeben. In diesem Voraussagebildsynthesizer 1603 wird ein Voraussagebild 117 synthetisiert und auf der Grundlage des anhand der Informationen 1604 bestimmten Rundungsverfahrens ausgegeben. Im Blockanpassungsabschnitt 116 in Fig. 1 gibt es keine Elemente, die mit 1602 und 1604 aus Fig. 16 äquivalent sind, und das Voraussagebild wird nur durch positives Runden synthetisiert. Weiterhin kann das im Blockanpassungsabschnitt bestimmte Rundungsverfahren 1605 ausgegeben werden, und diese Informationen können dann in den Bitstrom gemultipliziert und übertragen werden.

[0040] Ein Voraussagebildsynthesizer 1700 eines Bilddecodierers, der von einem Codierverfahren unter Verwendung von Mehrfachrundungsverfahren erzeugte Bitströme decodieren kann, ist in Fig. 17 dargestellt. Zahlen, die mit denen in anderen Zeichnungsbestandteilen identisch sind, geben die gleichen Teile an. Durch Ersetzen des Voraussagebildsynthesizers 211 aus Fig. 2 durch 1700 können Mehrfachrundungsverfahren verwendet werden. In der Rundungsverfahren-Bestimmungsvorrichtung 1701 wird das Rundungsverfahren bestimmt, das zur Voraussagebildsynthese beim Decodierprozeß geeignet ist. Zum richtigen Ausführen des Decodierens muß das hier gewählte Rundungsverfahren dem Rundungsverfahren gleichen, das zum Codieren gewählt wurde. Beispielsweise kann die folgende Regel zwischen dem Codierer und dem Decodierer geteilt werden: Wenn der aktuelle Rahmen ein P-Rahmen ist und die Anzahl der vom letzten I-Rahmen gezählten P-Rahmen (einschließlich des aktuellen Rahmens) ungerade ist, ist der aktuelle Rahmen ein (P+)-Rahmen. Wenn diese Zahl gerade ist, ist der aktuelle Rahmen ein (P-)-Rahmen. Falls die Rundungsverfahren-Bestimmungsvorrichtung auf der Codierseite (beispielsweise 1602 in Fig. 16) und die Rundungsverfahren-Bestimmungsvorrichtung 1701 mit dieser gemeinsamen Regel übereinstimmen, können die Bilder richtig decodiert werden. Das Voraussagebild wird unter Verwendung von Bewegungsinformationen 202, des Decoderbilds 210 des vorhergehenden Rahmens und von Informationen 1702, die sich auf das wie gerade beschrieben bestimmte Rundungsverfahren beziehen, im Voraussagebildsynthesizer 1703 synthetisiert. Dieses Voraussagebild 212 wird ausgegeben und dann zur Synthese des decodierten Bilds verwendet. Als eine Alternative zum vorstehend erwähnten Fall kann auch ein Fall erwogen werden, in dem die Informationen, die sich auf das Rundungsverfahren beziehen, in den übertragenen Bitstrom gemultipliziert sind (dieser Bitstrom kann am Codierer

DE 698 17 460 T2 2004.06.09

durch Ausgeben der Informationen 1605, die sich auf das Rundungsverfahren vom in Fig. 16 dargestellten Blockanpassungsabschnitt beziehen, erzeugt werden). In diesem Fall wird die Rundungsverfahren-Bestimmungsvorrichtung 1701 nicht verwendet, und die aus dem codierten Bitstrom entnommenen Informationen 1704, die sich auf das Rundungsverfahren beziehen, werden beim Voraussagebildsynthetisierer 1703 verwendet.

[0041] Abgesehen von dem Bildcodierer, bei dem die kundenspezifischen Schaltungen und Chips aus dem Stand der Technik verwendet werden, wie in Fig. 1 dargestellt ist, kann diese Erfindung auch auf Softwarebildcodierer angewendet werden, bei denen Vielzweckprozessoren verwendet werden. Ein Softwarebildcodierer 600 und ein Softwarebilddecodierer 700 sind in den Fig. 6 und 7 dargestellt. Im Softwarebildcodierer 600 wird ein Eingangsbild 601 zuerst im Eingangsrahmenspeicher 602 gespeichert, und der Vielzweckprozessor 603 lädt Informationen von hier und führt eine Codierung aus. Das Programm zum Ansteuern dieses Vielzweckprozessors wird von einer Speichervorrichtung 608 geladen, die eine Festplatte, eine Diskette usw. sein kann, und in einem Programmspeicher 604 gespeichert. Dieser Vielzweckprozessor verwendet auch einen Prozeßspeicher 605 zum Ausführen des Codierens. Die vom Vielzweckprozessor ausgegebenen Codierinformationen werden im Ausgangspuffer 606 zwischengespeichert und dann als ein codierter Bitstrom 607 ausgegeben.

[0042] Ein Flußdiagramm für die Codiersoftware (vom Computer lesbares Aufzeichnungsmedium) ist in Fig. 8 dargestellt. Der Prozeß beginnt bei 801, und der Wert 0 wird der Variablen N zugewiesen, wenn der Wert von N 100 ist. N ist ein Zähler für die Anzahl der Rahmen. 1 wird für jeden Rahmen addiert, dessen Verarbeitung abgeschlossen ist, und es sind Werte von 0 bis 99 erlaubt, wenn eine Codierung ausgeführt wird. Wenn der Wert für N 0 ist, ist der aktuelle Rahmen ein I-Rahmen. Wenn N eine ungerade Zahl ist, ist der aktuelle Rahmen ein (P+)-Rahmen, und wenn sie eine von 0 verschiedene gerade Zahl ist, ist der aktuelle Rahmen ein (P-)-Rahmen. Wenn die obere Grenze für den Wert von N 99 ist, bedeutet dies, daß ein I-Rahmen codiert wird, nachdem 99 P-Rahmen ((P+)-Rahmen oder (P-)-Rahmen) codiert worden sind. Durch stets erfolgreiches Einfügen eines Rahmens in eine bestimmte Anzahl codierter Rahmen können die folgenden Vorteile erhalten werden: (a) Eine Fehlerakkumulation infolge einer Nichtübereinstimmung zwischen der Verarbeitung des Codierers und derjenigen des Decodierers kann verhindert werden (beispielsweise eine Nichtübereinstimmung bei der Berechnung von DCT), und (b) die Verarbeitungsbelastung zum Gewinnen des reproduzierten Bilds des Zielrahmens anhand der codierten Daten (zufälliger Zugriff) wird verringert. Der optimale N-Wert ändert sich, wenn sich die Codierfunktionsweise oder die Umgebung, worin der Codierer verwendet wird, ändert. Dies bedeutet daher nicht, daß der Wert von N stets 100 sein muß. Der Prozeß zum Bestimmen des Rundungsverfahrens und des Codiermodus für jeden Rahmen wird bei 805 ausgeführt, und das Flußdiagramm mit Einzelheiten dieser Operation ist in Fig. 9 dargestellt. Zuerst wird bei 901 geprüft, ob N eine 0 ist oder nicht. Falls N 0 ist, wird "I" als Unterscheidungsinformationen des Voraussagemodus bei 902 an den Ausgangspuffer ausgegeben. Dies bedeutet, daß das zu codierende Bild als ein I-Rahmen codiert wird. Hierbei bedeutet "Ausgabe an den Ausgangspuffer", daß die Informationen nach dem Speichern im Ausgangspuffer als ein Teil des codierten Bitstroms an eine externe Vorrichtung ausgegeben werden. Wenn N nicht 0 ist, wird bei 904 festgestellt, ob N eine ungerade oder eine gerade Zahl ist. Wenn N eine ungerade Zahl ist, wird "+" bei 905 als Unterscheidungsinformationen für das Rundungsverfahren an den Ausgangspuffer ausgegeben, und das zu codierende Bild wird als ein (P+)-Rahmen codiert. Wenn N andererseits eine gerade Zahl ist, wird "-" bei 906 als Unterscheidungsinformationen für das Rundungsverfahren an den Ausgangspuffer ausgegeben, und das zu codierende Bild wird als ein (P-)-Rahmen codiert. Der Prozeß kehrt wieder zu Fig. 8 zurück, wo nach dem Bestimmen des Codiermodus bei 805 das Eingangsbild bei 806 im Rahmenspeicher A gespeichert wird. Der hier erwähnte Rahmenspeicher A bezeichnet einen Abschnitt der Speicherzone (beispielsweise die im Speicher von 605 in Fig. 6 unterhaltene Speicherzone) des Softwarecodierers. Es wird bei 807 geprüft, ob der gegenwärtig codierte Rahmen ein I-Rahmen ist. Wenn er nicht als ein I-Rahmen identifiziert wird, werden bei 808 eine Bewegungsschätzung und eine Bewegungskompensation ausgeführt. Das Flußdiagramm in Fig. 10 zeigt Einzelheiten dieses bei 808 ausgeführten Prozesses. Zuerst wird bei 1001 eine Bewegungsschätzung zwischen den in den Rahmenspeichern A und B gespeicherten Bildern ausgeführt (wie im letzten Teil dieses Abschnitts geschrieben ist, ist das decodierte Bild des vorhergehenden Rahmens im Rahmenspeicher B gespeichert). Es wird der Bewegungsvektor für jeden Block gefunden, und dieser Bewegungsvektor wird zum Ausgangspuffer gesendet. Als nächstes wird bei 1002 geprüft, ob der aktuelle Rahmen ein (P+)-Rahmen ist. Wenn der aktuelle Rahmen ein (P+)-Rahmen ist, wird das Voraussagebild bei 1003 unter Verwendung des positiven Rundens synthetisiert, und dieses Voraussagebild wird im Rahmenspeicher C gespeichert. Wenn der aktuelle Rahmen andererseits ein (P-)-Rahmen ist, wird das Voraussagebild bei 1004 unter Verwendung des negativen Rundens synthetisiert, und dieses Voraussagebild wird im Rahmenspeicher C gespeichert. Als nächstes wird bei 1005 das differentielle Bild zwischen den Rahmenspeichern A und C gefunden und im Rahmenspeicher A gespeichert. Hier kehrt der Prozeß wieder zu Fig. 8 zurück. Vor Beginn der Verarbeitung bei 809 wird das Eingangsbild im Rahmenspeicher A gespeichert, wenn der aktuelle Rahmen ein I-Rahmen ist, und das differentielle Bild zwischen dem Eingangsbild und dem Voraussagebild wird im Rahmenspeicher A gespeichert, wenn



DE 698 17 460 T2 2004.06.09

der aktuelle Rahmen ein P-Rahmen ((P+)-Rahmen oder (P-)-Rahmen) ist. Bei 809 wird DCT auf das im Rahmenspeicher A gespeicherte Bild angewendet, und die hier berechneten DCT-Koeffizienten werden nach dem Quantisieren zum Ausgangspuffer gesendet. Bei 810 wird eine inverse Quantisierung an den quantisierten DCT-Koeffizienten ausgeführt, und es wird eine inverse DCT angewendet. Das durch Anwenden der inversen DCT erhaltene Bild wird im Rahmenspeicher B gespeichert. Als nächstes wird bei 811 wieder geprüft, ob der aktuelle Rahmen ein I-Rahmen ist. Wenn der aktuelle Rahmen kein I-Rahmen ist, werden die in den Rahmenspeichern B und C gespeicherten Bilder addiert, und das Ergebnis wird im Rahmenspeicher B gespeichert. Der Prozeß zum Codieren eines Rahmens endet hier, und das im Rahmenspeicher B gespeicherte Bild, bevor zu 813 übergegangen wird, ist das rekonstruierte Bild dieses Rahmens (dieses Bild ist mit dem an der Decodierseite erhaltenen identisch). Bei 813 wird geprüft, ob der Rahmen, dessen Codierung gerade abgeschlossen wurde, der letzte Rahmen in der Folge ist. Falls dies der Fall ist, endet der Codierprozeß. Falls dieser Rahmen nicht der letzte Rahmen ist, wird 1 bei 814 zu N addiert, und der Prozeß kehrt wieder zu 803 zurück, und es beginnt der Codierprozeß für den nächsten Rahmen.

[0043] Ein Softwaredecoder 700 ist in Fig. 7 dargestellt. Nachdem der codierte Bitstrom 701 im Eingangspuffer 702 zwischengespeichert wurde, wird dieser Bitstrom in den Vielzweckprozessor 703 geladen. Das Programm zum Ansteuern dieses Vielzweckprozessors wird von einer Speichervorrichtung 708 geladen, die eine Festplatte, eine Diskette usw. sein kann, und in einem Programmspeicher 704 gespeichert. Dieser Vielzweckprozessor verwendet auch einen Prozeßspeicher 605 zum Ausführen des Decodierens. Das durch den Decodierprozeß erhaltene decodierte Bild wird im Ausgangsrahmenspeicher 706 zwischengespeichert und dann als Ausgangsbild 707 nach außen gesendet.

[0044] Ein Flußdiagramm der Decodersoftware für den in Fig. 7 dargestellten Softwaredecoder 700 ist in Fig. 11 dargestellt. Der Prozeß beginnt bei 1101, und es wird bei 1102 geprüft, ob Eingangsinformationen vorhanden sind. Falls keine Eingangsinformationen vorhanden sind, endet der Decodierprozeß bei 1103. Wenn Eingangsinformationen vorhanden sind, werden Unterscheidungsinformationen des Voraussagemodus bei 1104 eingegeben. Das hier verwendete Wort "Eingang" bedeutet, daß die im Eingangspuffer (beispielsweise 702 von Fig. 7) gespeicherten Informationen durch den Vielzweckprozessor geladen werden. Bei 1105 wird geprüft, ob die Codiermodus-Unterscheidungsinformationen "I" sind. Wenn dies nicht der Fall ist, werden die Unterscheidungsinformationen für das Rundungsverfahren eingegeben, und die Synthese des Zwischenrahmen-Voraussagebilds wird bei 1107 ausgeführt. Ein Flußdiagramm, in dem Einzelheiten des bei 1107 ausgeführten Vorgangs dargestellt sind, ist in Fig. 12 dargestellt. Bei 1201 wird ein Bewegungsvektor für jeden Block eingegeben. Daraufhin wird bei 1202 geprüft, ob die bei 1106 geladenen Unterscheidungsinformationen für das Rundungsverfahren ein "+" sind. Wenn dies der Fall ist, ist der gerade decodierte Rahmen ein (P+)-Rahmen. In diesem Fall wird das Voraussagebild unter Verwendung eines positiven Rundens bei 1203 synthetisiert, und das Voraussagebild wird im Rahmenspeicher D gespeichert. Hier bezeichnet der Rahmenspeicher

D einen Abschnitt der Speicherzone des Softwaredecoders (diese Speicherzone wird beispielsweise im Verarbeitungsspeicher 705 in Fig. 7 erhalten).

Wenn die Unterscheidungsinformationen des Rundungsverfahrens

ein "-" sind, ist der gerade decodierte aktuelle Rahmen ein (P-)-Rahmen. Das Voraussagebild wird bei 1204

unter Verwendung des negativen Rundens synthetisiert, und dieses Voraussagebild wird im Rahmenspeicher

D gespeichert. Falls an diesem Punkt ein (P+)-Rahmen infolge irgendeines Fehlers als ein (P-)-Rahmen de-

codiert wird oder falls umgekehrt ein (P-)-Rahmen als ein (P+)-Rahmen decodiert wird, wird das richtige Vor-

sagebild nicht im Decoder synthetisiert, und die Qualität des decodierten Bilds nimmt ab. Nach dem Syn-

thetisieren des Voraussagebilds kehrt der Vorgang zu Fig. 11 zurück, und die quantisierten DCT-Koeffizienten

werden bei 1108 eingegeben. Es werden dann eine inverse Quantisierung und eine inverse DCT auf diese Ko-

effizienten angewendet, und das sich ergebende Bild wird im Rahmenspeicher E gespeichert. Bei 1109 wird

wieder geprüft, ob der gerade decodierte Rahmen ein I-Rahmen ist. Falls der aktuelle Rahmen kein I-Rahmen

ist, werden die in den Rahmenspeichern D und E gespeicherten Bilder bei 1110 addiert, und das sich ergebende

Summenbild wird im Rahmenspeicher E gespeichert. Das im Rahmenspeicher E vor Beginn des Prozesses

bei 1111 gespeicherte Bild ist das rekonstruierte Bild. Dieses im Rahmenspeicher E gespeicherte Bild wird bei

1111 an den Ausgangsrahmenspeicher (beispielsweise 706 in Fig. 7) ausgegeben und dann vom Decoder

als das rekonstruierte Bild ausgegeben. Der Decodierprozeß für einen Rahmen wird hier abgeschlossen, und

der Prozeß für den nächsten Rahmen beginnt mit der Rückkehr zu 1102.

[0045] Wenn eine auf dem in den Fig. 6-12 dargestellten Flußdiagramm beruhende Software in den Soft-

warebildcodierern oder -decodierern abläuft, wird die gleiche Wirkung erhalten, als ob kundenspezifische

Schaltungen und kundenspezifische Chips verwendet werden.

[0046] Ein Speichermedium (Aufzeichnungsmedium), bei dem der vom Softwarecodierer 601 aus Fig. 6 er-

zeugte Bitstrom aufgezeichnet wird, ist in Fig. 13 dargestellt. Es wird angenommen, daß die in den Flußdia-

grammen aus den Fig. 8-10 dargestellten Algorithmen im Softwarecodierer verwendet werden. Digitale Infor-

mationen werden konzentrisch auf einer Aufzeichnungsplatte 1301 aufgezeichnet, die zur Aufzeichnung digi-

taler Informationen in der Lage ist (wobei es sich beispielsweise um magnetische Platten, eine optische Platte

usw. handelt). Ein Teil 1302 der auf dieser digitalen Platte aufgezeichneten Informationen umfaßt: Voraussa-

DE 698 17 460 T2 2004.06.09

gemodus-Unterscheidungsinformationen 1303, 1305, 1308, 1311 und, 1314, Rundungsverfahren-Unterscheidungsinformationen 1306, 1309, 1312 und 1315 und Bewegungsvektor- und DCT-Koeffizienteninformationen 1304, 1307, 1310, 1313 und 1316. "I" darstellende Informationen werden bei 1303 aufgezeichnet, "P" darstellende Informationen werden bei 1305, 1308, 1311 und 1314 aufgezeichnet, "+" darstellende Informationen werden bei 1309 und 1315 aufgezeichnet. In diesem Fall können "I" und "+" durch ein einziges Bit 0 dargestellt werden, und "P" und "-" können durch ein einziges Bit 1 dargestellt werden. Unter Verwendung dieser Darstellung kann der Decodierer die aufgezeichneten Informationen richtig interpretieren, und das korrigierte rekonstruierte Bild wird synthetisiert. Durch Speichern eines codierten Bitstroms in einem Speichermedium unter Verwendung des vorstehend beschriebenen Verfahrens wird die Akkumulation von Rundungsfehlern verhindert, wenn der Bitstrom gelesen und decodiert wird.

[0047] Ein Speichermedium, bei dem der Bitstrom der codierten Daten der in Fig. 5 dargestellten Bildfolge aufgezeichnet wird, ist in Fig. 15 dargestellt. Der aufgezeichnete Bitstrom enthält Informationen, die sich auf (P+)-, (P-)- und B-Rahmen beziehen. Ebenso wie bei 1301 aus Fig. 13 werden digitale Informationen konzentrisch auf einer Aufzeichnungsplatte 1501 aufgezeichnet, welche zum Aufzeichnen digitaler Informationen in der Lage ist (wobei es sich beispielsweise um magnetische Platten, optische Platten usw. handelt). Ein Teil 1502 der auf dieser digitalen Platte aufgezeichneten Informationen umfaßt: Voraussagemodus-Unterscheidungsinformationen 1503, 1505, 1508, 1510 und 1513, Rundungsverfahren-Unterscheidungsinformationen 1506 und 1512 und Bewegungsvektor- und DCT-Koeffizienteninformationen 1504, 1507, 1509, 1511 und 1514. "I" darstellende Informationen werden bei 1503 aufgezeichnet, "P" darstellende Informationen werden bei 1505 und 1510 aufgezeichnet, "B" darstellende Informationen werden bei 1508 und 1513 aufgezeichnet, "+" darstellende Informationen werden bei 1505 aufgezeichnet, und "-" darstellende Informationen werden bei 1511 aufgezeichnet. In diesem Fall können "I", "P" und "B" jeweils durch Zweibitwerte 00, 01 und 10 dargestellt werden, und "+" und "-" können jeweils durch Einbitwerte 0 und 1 dargestellt werden. Unter Verwendung dieser Darstellung kann der Decodierer die aufgezeichneten Informationen richtig interpretieren, und das korrigierte rekonstruierte Bild wird synthetisiert. In Fig. 15 sind auf den Rahmen 501 (I-Rahmen) in Fig. 5 bezogene Informationen 1503 und 1504, auf den Rahmen 502 (B-Rahmen) bezogene Informationen 1508 und 1509, auf den Rahmen 503 ((P+)-Rahmen) bezogene Informationen 1505 und 1507, auf den Rahmen 504 (B-Rahmen) bezogene Informationen 1513 und 1514 und auf den Rahmen 505 ((P-)-Rahmen) bezogene Informationen 1510 und 1512. Wenn Codierbildfolgen unter Verwendung von B-Rahmen codiert werden, ist die Übertragungsreihenfolge und die Anzeigereihenfolge der Rahmen gewöhnlich verschieden. Dies liegt daran, daß die vorausgegangenen und die folgenden Referenzbilder codiert werden müssen, bevor das Voraussagebild für den B-Rahmen synthetisiert wird. Folglich werden Informationen, die sich auf den Rahmen 503 beziehen, trotz der Tatsache, daß der Rahmen 502 vor dem Rahmen 503 angezeigt wird, vor Informationen übertragen, die sich auf den Rahmen 502 beziehen. Wie vorstehend beschrieben wurde, ist es nicht erforderlich, Mehrfachrundungsverfahren für B-Rahmen zu verwenden, weil die Bewegungskompensation bei B-Rahmen keine Akkumulation von Rundungsfehlern hervorruft. Wie in diesem Beispiel gezeigt wurde, werden daher Informationen, die Rundungsverfahren (beispielsweise "+" und "-") bezeichnen, für B-Rahmen nicht übertragen. Selbst wenn nur ein positives Runden auf B-Rahmen angewendet wird, tritt das Problem akkumulierter Rundungsfehler daher beispielsweise nicht auf. Durch Speichern codierter Bitströme, welche auf B-Rahmen bezogene Informationen enthalten, in der vorstehend beschriebenen Weise in einem Speichermedium kann das Auftreten akkumulierter Rundungsfehler verhindert werden, wenn dieser Bitstrom gelesen und decodiert wird.

[0048] Spezifische Beispiele von Codierern und Decodierern unter Verwendung des in dieser Beschreibung dargelegten Codierverfahrens sind in Fig. 14 dargestellt. Das Bildcodier- und -decodierverfahren kann durch Installieren einer Bildcodier- und -decodiersoftware in einem Computer 1401 verwirklicht werden. Diese Software ist auf irgendeiner Art eines Speichermediums (CDROM, Diskette, Festplatte usw.) 1412 aufgezeichnet, wird in einen Computer geladen und dann verwendet. Zusätzlich kann der Computer als ein Bildkommunikations-Endgerät verwendet werden, indem der Computer an Kommunikationsleitungen angeschlossen wird. Es ist auch möglich, das in dieser Beschreibung dargelegte Decodierverfahren in einer Abspielvorrichtung 1403 zu installieren, die den auf einem Speichermedium 1402 aufgezeichneten codierten Bitstrom liest und decodiert. In diesem Fall kann das rekonstruierte Bildsignal auf einem Fernsehbildschirm 1404 angezeigt werden. Die Vorrichtung 1403 kann nur zum Lesen des codierten Bitstroms verwendet werden, und die Decodiervorrichtung kann in diesem Fall im Fernsehbildschirm 1404 installiert werden. Es ist wohlbekannt, daß eine digitale Datenübertragung unter Verwendung von Satelliten und Erdwellen verwirklicht werden kann. Eine Decodiervorrichtung kann auch in einem Fernsehempfänger 1405 installiert werden, der in der Lage ist, diese digitalen Übertragungen zu empfangen. Weiterhin kann eine Decodiervorrichtung auch in einer Settop-Box 1409 eingerichtet werden, die an eine Satelliten-/Erdwellenantenne oder ein Kabel 1408 eines Kabelfernsehsystems angeschlossen ist, so daß die rekonstruierten Bilder auf einem Fernsehbildschirm 1410 angezeigt werden können. In diesem Fall kann die Decodiervorrichtung im Fernsehbildschirm statt in der Settop-Box untergebracht werden, wie es bei 1404 der Fall ist. Das Layout eines digitalen Satellitenausstrahlungssystems ist bei 1413,

DE 698 17 460 T2 2004.06.09

1414 und 1415 dargestellt. Die Videoinformationen in dem codierten Bitstrom werden von einer Ausstrahlungsstation 1413 zu einem Kommunikations- oder Ausstrahlungssatelliten 1414 übertragen. Der Satellit empfängt diese Informationen, sendet sie zu einem Heim 1415 mit einer Ausrüstung zum Empfangen von Satellitenausstrahlungsprogrammen, und die Videoinformationen werden in diesem Heim unter Verwendung von Vorrichtungen, wie einem Fernsehempfänger oder einer Settop-Box rekonstruiert und angezeigt. Die digitale Bildkommunikation unter Verwendung mobiler Endgeräte 1406 hat kürzlich erhebliches Interesse gefunden, weil diese Bildkommunikation bei sehr niedrigen Bitraten möglich geworden ist. Digitale tragbare Endgeräte können in die folgenden drei Typen eingeteilt werden, nämlich einen Transceiver mit einem Codierer und einem Decodierer, einen Sender, der nur einen Codierer aufweist, und einen Empfänger, der nur einen Decodierer aufweist. Eine Codiervorrichtung kann in einem Videokamerarecorder 1407 installiert werden. Die Kamera kann auch lediglich zum Aufnehmen des Bildsignals verwendet werden, und dieses Signal kann einem kundenspezifischen Codierer 1411 zugeführt werden. Alle in dieser Zeichnung dargestellten Vorrichtungen oder Systeme können mit dem in dieser Beschreibung dargestellten Codier- und/oder Decodierverfahren ausgestattet werden. Bei Verwendung dieses Codier- und/oder Decodierverfahrens in diesen Vorrichtungen oder Systemen können Bilder erhalten werden, die eine höhere Qualität aufweisen als jene, die unter Verwendung herkömmlicher Technologien erhalten werden.

[0049] Die folgenden Abänderungen sind klar im Schutzzumfang dieser Erfindung enthalten.

(i) Eine Vorbedingung des vorstehend beschriebenen Prinzips war die Verwendung einer Blockanpassung als ein Bewegungskompensationsverfahren. Diese Erfindung kann jedoch weiter auch auf alle Bildfolge-Decodierverfahren angewendet werden, bei denen eine Bewegungskompensation unter Verwendung eines Werts für die vertikalen und horizontalen Komponenten des Pixelbewegungsvektors ausgeführt wird, welcher sich von einem ganzzahligen Vielfachen der Abtastperiode der Pixel in vertikaler und horizontaler Richtung unterscheidet, und indem dann durch Interpolation der Intensitätswert einer Position gefunden wird, an der der Abtastwert nicht vorhanden ist. Demgemäß sind beispielsweise die in der japanischen Patentanmeldung Hei 08-060572 angeführte globale Bewegungskompensation und die in der japanischen Patentanmeldung Hei 08-249601 angeführte Warpvorhersage auf das Verfahren gemäß dieser Erfindung anwendbar.

(ii) In der Beschreibung der Erfindung wurde nur der Fall erwähnt, in dem ein ganzzahliges Vielfaches von  $1/2$  für die horizontalen und vertikalen Komponenten des Bewegungsvektors verwendet wurde. Diese Erfindung ist jedoch allgemein auch auf Verfahren anwendbar, bei denen ganzzahlige Vielfache von  $1/d$  ( $d$  ist eine positive ganze Zahl und auch eine gerade Zahl) für die horizontalen und vertikalen Komponenten des Bewegungsvektors zulässig sind. Wenn  $d$  jedoch groß wird, wird auch der Teiler für die Division bei der bilinearen Interpolation (Quadrat von  $d$ , siehe Gleichung 2) groß, so daß dagegen die Wahrscheinlichkeit, daß Ergebnisse einer normalen Division einen Wert von 0,5 erreichen, gering wird. Wenn dementsprechend nur ein positives Runden ausgeführt wird, wird der Absolutwert der Erwartung von Rundungsfehlern klein, und die durch akkumulierte Fehler hervorgerufenen nachteiligen Wirkungen werden weniger erheblich. Auf das Verfahren gemäß dieser Erfindung ist weiterhin ein Bewegungskompensationsverfahren anwendbar, bei dem der Wert beispielsweise veränderlich ist und sowohl ein positives als auch ein negatives Runden verwendet wird, wenn  $d$  kleiner als ein fester Wert ist, und nur ein positives Runden oder nur ein negatives Runden verwendet wird, wenn der Wert von  $d$  größer als ein fester Wert ist.

(iii) Wie im Stand der Technik erwähnt wurde, treten die nachteiligen Wirkungen von akkumulierten Rundungsfehlern leicht auf, wenn die quantisierte Schrittgröße der DCT-Koeffizienten groß ist, sofern DCT als ein Fehlercodierverfahren verwendet wird. Es ist jedoch auch ein Verfahren auf die Erfindung anwendbar, bei dem sowohl ein positives als auch ein negatives Runden verwendet wird, wenn die Quantisierungsschrittgröße von DCT-Koeffizienten größer als ein Schwellenwert ist. Wenn die Quantisierungsschrittgröße der DCT-Koeffizienten kleiner als der Schwellenwert ist, wird nur ein positives Runden oder nur ein negatives Runden verwendet.

(iv) In Fällen, in denen in der Luminanzebene Fehlerakkumulationen auftreten und in Fällen, in denen in der Chrominanzebene Fehlerakkumulationen auftreten, sind die nachteiligen Wirkungen auf die rekonstruierten Bilder im allgemeinen im Fall von Fehlerakkumulationen in der Chrominanzebene ernster. Dies liegt daran, daß Fälle, in denen Gesamtänderungen der Bildfarbe auftreten, auffälliger sind als Fälle, in denen sich das Bild verdunkelt oder leicht aufhellt. Es ist jedoch auch ein Verfahren auf diese Erfindung anwendbar, bei dem sowohl ein positives als auch ein negatives Runden für das Chrominanzsignal verwendet wird und nur ein positives Runden oder ein negatives Runden für das Luminanzsignal verwendet wird.

[0050] Wie in der Beschreibung des Stands der Technik beschrieben wurde, werden durch Halbieren der  $1/2$ -Pixelgenauigkeits-Bewegungsvektoren erhaltene  $1/4$ -Pixelgenauigkeits-Bewegungsvektoren bei H.263 zu einer  $1/2$ -Pixelgenauigkeit gerundet. Durch Hinzufügen bestimmter Änderungen zu diesem Verfahren kann jedoch der absolute Erwartungswert für Rundungsfehler verringert werden. Bei H.263, welches im Stand der Technik erwähnt wurde, wird ein Wert, der die Hälfte der horizontalen oder vertikalen Komponenten des Be-

## DE 698 17 460 T2 2004.06.09

wegungsvektors für die Luminanzebene beträgt, als  $r + s/4$  ausgedrückt ( $r$  ist eine ganze Zahl,  $s$  ist eine ganze Zahl kleiner als 4 und nicht kleiner als 0), und wenn  $s$  1 oder 3 ist, wird eine Rundungsoperation ausgeführt, um eine 2 zu erhalten. Dieser Vorgang kann folgendermaßen geändert werden: Wenn  $s$  1 ist, wird eine Rundungsoperation ausgeführt, um eine 0 zu erhalten, und wenn  $s$  3 ist, wird 1 zu  $r$  addiert, um  $s$  zu 0 zu machen. Durch Ausführen dieser Operationen wird die Häufigkeit, mit der die Intensitätswerte an den Positionen 406-408 in Fig. 4 berechnet werden (die Wahrscheinlichkeit, daß horizontale und vertikale Komponenten des Bewegungsvektors zu einer großen ganzen Zahl werden), erheblich verringert, so daß der absolute Erwartungswert für den Rundungsfehler klein wird. Selbst dann, wenn die Größe des bei diesem Verfahren auftretenden Fehlers begrenzt werden kann, kann die Akkumulation von Fehlern jedoch nicht vollständig verhindert werden.

(v) Die in dieser Beschreibung dargelegte Erfindung ist auf ein Verfahren anwendbar, das das endgültige Zwischenrahmen-Voraussagebild durch Mitteln der durch verschiedene Bewegungskompensationsverfahren erhaltenen Voraussagebilder erhält. In dem in der japanischen Patentanmeldung Hei 8-2816 beschriebenen Verfahren werden beispielsweise durch die folgenden zwei Verfahren erhaltene Zwischenrahmen-Voraussagebilder gemittelt: eine Blockanpassung, bei der ein Bewegungsvektor jedem  $16 \times 16$ -Pixelblock zugewiesen wird, und eine Blockanpassung; bei der ein Bewegungsvektor allen  $8 \times 8$ -Pixelblöcken zugewiesen wird. Bei diesem Verfahren wird das Runden auch ausgeführt, wenn der Mittelwert der zwei Voraussagebilder berechnet wird. Wenn bei dieser Mittelungsoperation nur ein positives Runden kontinuierlich ausgeführt wird, wird ein neuer Typ eines Rundungsfehlers akkumuliert. Dieses Problem kann unter Verwendung von Mehrfachrundungsverfahren für diese Mittelungsoperation gelöst werden. Bei diesem Verfahren wird bei der Mittelungsoperation ein negatives Runden ausgeführt, wenn ein positives Runden bei der Blockanpassung ausgeführt wird. Umgekehrt wird das positive Runden für das Mitteln verwendet, wenn das negative Runden für die Blockanpassung verwendet wird. Unter Verwendung verschiedener Rundungsverfahren für das Mitteln und die Blockanpassung werden die Rundungsfehler von zwei verschiedenen Quellen innerhalb des gleichen Rahmens aufgehoben.

(vi) Wenn ein Verfahren verwendet wird, bei dem alternierend (P+)-Rahmen und (P-)-Rahmen entlang der Zeitachse lokalisiert werden, muß der Decodierer bestimmen, ob der gegenwärtig verarbeitete P-Rahmen ein (P+)-Rahmen oder ein (P-)-Rahmen ist. Nachstehend wird ein Beispiel eines solchen Identifikationsverfahrens angegeben: Ein Zähler zählt die Anzahl der P-Rahmen nach dem letzten codierten oder decodierten I-Rahmen, und der aktuelle P-Rahmen ist ein (P+)-Rahmen, wenn die Zahl ungerade ist, und ein (P-)-Rahmen, wenn die Zahl gerade ist (dieses Verfahren wird als ein implizites Schema bezeichnet). Es gibt beispielsweise auch ein Verfahren, das in den Kopfabschnitt der codierten Bildinformationen Informationen schreibt, die der Angabe dienen, ob der gegenwärtig codierte P-Rahmen am Codierer ein (P+)-Rahmen oder ein (P-)-Rahmen ist (dieses Verfahren wird als ein explizites Schema bezeichnet). Verglichen mit dem impliziten Verfahren, ist dieses Verfahren gut in der Lage, Übertragungsfehler zu widerstehen, weil es nicht erforderlich ist, die Anzahl der P-Rahmen zu zählen.

[0051] Weiterhin hat das explizite Verfahren die folgenden Vorteile: Wie in Beschreibung des Stands der Technik beschrieben wurde, verwenden frühere Codierstandards (wie MPEG-1 oder MPEG-2) nur ein positives Runden zur Bewegungskompensation. Dies bedeutet beispielsweise, daß die Bewegungsschätz- bzw. Bewegungskompensationsvorrichtungen (die beispielsweise zu 106 in Fig. 1 äquivalent sind) für MPEG-1/MPEG-2, die auf dem Markt erhältlich sind, nicht mit Codierverfahren kompatibel sind, welche sowohl (P+)-Rahmen als auch (P-)-Rahmen verwenden. Es wird angenommen, daß es einen Decodierer gibt, der Bitströme decodieren kann, welche von einem Codierverfahren erzeugt wurden, das (P+)-Rahmen und (P-)-Rahmen verwendet. Falls in diesem Fall der Decodierer auf dem vorstehend erwähnten impliziten Verfahren beruht, ist es schwierig, einen Codierer zu entwickeln, der Bitströme erzeugt, welche vom vorstehend erwähnten Decodierer unter Verwendung der vorstehend erwähnten Bewegungsschätz- und -kompensationsvorrichtung für MPEG-1/MPEG-2 richtig decodiert werden können. Falls der Decodierer jedoch auf dem vorstehend erwähnten expliziten Verfahren beruht, kann dieses Problem gelöst werden. Ein Codierer, bei dem eine MPEG-1/MPEG-2-Bewegungsschätz- und -kompensationsvorrichtung verwendet wird, kann durch kontinuierliches Schreiben von Rundungsverfahren-Unterscheidungsinformationen, die ein positives Runden angeben, in den Rahmeninformations-Kopfteil kontinuierlich (P+)-Rahmen senden. Wenn dies erfolgt, kann ein auf dem expliziten Verfahren beruhender Decodierer den von diesem Codierer erzeugten Bitstrom richtig decodieren. Es sollte in diesem Fall natürlich wahrscheinlicher sein, daß die Akkumulation von Rundungsfehlern auftritt, weil nur (P+)-Rahmen vorhanden sind. Die Fehlerakkumulation ist jedoch in Fällen, in denen der Codierer nur kleine Werte als Quantisierungsschrittgröße für die DCT-Koeffizienten verwendet, kein ernstes Problem (ein Beispiel für solche Codierer ist ein kundenspezifischer Codierer, der nur zum Codieren bei einer hohen Rate verwendet wird). Zusätzlich zu dieser Austauschbarkeit zwischen früheren Standards hat das explizite Verfahren weiter die folgenden Vorteile: (a) die Gerätekosten für bei einer hohen Rate arbeitende kundenspezifische Codierer und Decodierer, die nicht für eine Rundungsfehlerakkumulation infolge des häufigen Einführens von

DE 698 17 460 T2 2004.06.09

I-Rahmen anfällig sind, kann durch Installieren nur eines positiven oder eines negativen Rundens als das Pixelwert-Rundungsverfahren zur Bewegungskompensation verringert werden, und (b) die erwähnten Codierer, die nicht für eine Rundungsfehlerakkumulation anfällig sind, haben den Vorteil, daß es nicht erforderlich ist, zu entscheiden, ob der aktuelle Rahmen als ein (P+)- oder ein (P-)-Rahmen codiert werden muß, und daß die Verarbeitung vereinfacht ist.

(vii) Die in dieser Beschreibung dargelegte Erfindung ist auf Decodierverfahren anwendbar, die ein das Runden begleitendes Filtern auf die Zwischenrahmen-Voraussagebilder anwenden. Beim internationalen Standard H.261 für die Bildfolgecodierung wird beispielsweise ein Tiefpaßfilter (als ein Schleifenfilter bezeichnet) auf Blocksignale angewendet, deren Bewegungsvektoren bei Zwischenrahmen-Voraussagebildern nicht 0 sind. Weiterhin können bei H.263 Filter zum Herausglätten von Unstetigkeiten an Blockgrenzen (Blockartefakte) verwendet werden. All diese Filter führen ein gewichtetes Mittel von Pixelintensitätswerten aus, und es wird dann an den gemittelten Intensitätswerten ein Runden ausgeführt. Selbst in diesen Fällen ist die selektive Verwendung eines positiven Rundens und eines negativen Rundens zum Verhindern der Fehlerakkumulation wirksam.

(viii) Abgesehen von  $P+ P- P+ P- \dots$  sind verschiedene Verfahren zum Mischen von (P+)-Rahmen und (P-)-Rahmen, wie  $P+ P+ P- P- P+ P+ \dots$  oder  $P+ P- P- P+ P+ \dots$  auf das Verfahren gemäß dieser Erfindung anwendbar. Bei Verwendung eines Zufallszahlengenerators, der 0 und 1 bei einer Wahrscheinlichkeit von 50 Prozent ausgibt, kann der Codierer beispielsweise einen (P+)- und einen (P-)-Rahmen codieren, wenn die Ausgabe 0 bzw. 1 ist. In jedem Fall ist es umso weniger wahrscheinlich, daß die Rundungsfehlerakkumulation auftritt, je kleiner die Wahrscheinlichkeitsdifferenz ist, daß (P+)-Rahmen und (P-)-Rahmen in einem bestimmten Zeitraum auftreten. Wenn der Codierer weiterhin (P+)-Rahmen und (P-)-Rahmen durch ein beliebiges Verfahren mischen darf, müssen der Codierer und der Decodierer auf der Grundlage des expliziten Verfahrens und nicht mit dem vorstehend beschriebenen impliziten Verfahren arbeiten. Dementsprechend ist das explizite Verfahren überlegen, wenn es aus der Perspektive des Ermöglichens einer Flexibilität Konfiguration für den Codierer und den Decodierer betrachtet wird.

(ix) Die in dieser Beschreibung dargelegte Erfindung begrenzt das Pixelwert-Interpolationsverfahren nicht auf die bilineare Interpolation. Interpolationsverfahren für Intensitätswerte können im allgemeinen durch die folgende Gleichung beschrieben werden:

[Gleichung 5]

$$R(x + r, y + s) = T \left( \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} h(x - j, s - k) R(x + j, y + k) \right) \quad (5)$$

[0052] Hierbei sind  $r$  und  $s$  reelle Zahlen, ist  $h(r, s)$  eine Funktion zum Interpolieren der reellen Zahlen und ist  $T(z)$  eine Funktion zum Runden der reellen Zahl  $z$ . Die Definitionen von  $R(x, y)$ ,  $x$  und  $y$  gleichen denen in Gleichung 4. Eine Bewegungskompensation unter Verwendung des positiven Rundens wird ausgeführt, wenn  $T(z)$  eine das positive Runden darstellende Funktion ist, und eine Bewegungskompensation unter Verwendung des negativen Rundens wird ausgeführt, wenn die Funktion das negative Runden darstellt. Diese Erfindung ist auf Interpolationsverfahren anwendbar, die unter Verwendung von Gleichung 5 beschrieben werden können. Beispielsweise kann eine bilineare Interpolation durch Definieren von  $h(r, s)$  beschrieben werden, wie nachstehend dargestellt ist.

[Gleichung 6]

$$h(r, s) = \begin{cases} (1 - |r|)(1 - |s|), & 0 \leq |r| \leq 1, 0 \leq |s| \leq 1 \\ 0, & \text{andernfalls} \end{cases} \quad (6)$$

[0053] Falls  $h(r, s)$  jedoch beispielsweise wie nachstehend angegeben definiert ist,

[Gleichung 7]

$$h(r, s) = \begin{cases} 1 - |r| - |s|, & 0 \leq |r| + |s| \leq 1, rs < 0, \\ 1 - |r|, & |r| \geq |s|, |r| \leq 1, rs \geq 0, \\ 1 - |s|, & |s| > |r|, |s| \leq 1, rs > 0, \\ 0, & \text{andernfalls,} \end{cases} \quad (7)$$

DE 698 17 460 T2 2004.06.09

wird ein Interpolationsverfahren implementiert, das von der bilinearen Interpolation verschieden ist, die Erfindung ist jedoch weiterhin anwendbar.

(x) Die in dieser Beschreibung dargelegte Erfindung beschränkt das Codierverfahren für Fehlerbilder nicht auf DCT (diskrete Kosinustransformation). Beispielsweise sind eine Wellenform-Transformation (beispielsweise M. Antonioni u. a., "Image Coding Using Wavelet Transform", IEEE Trans. Image Processing, Band 1, Nr. 2, April 1992) und eine Walsh-Hadamard-Transformation (beispielsweise A. N. Netravalli und B. G. Haskell, "Digital Pictures", Plenum Press, 1998) auch auf diese Erfindung anwendbar.

### Patentansprüche

#### 1. Verfahren zum Dekodieren von Bildern, wobei:

ein zuvor dekodiertes Bild als Referenzbild (210) gespeichert wird;  
Bewegungsvektordinformation (202) aus zu dekodierender Eingangsinformation gewonnen wird; und  
Bewegungskompensation unter Verwendung der Bewegungsvektordinformation und des Referenzbilds durchgeführt wird; um ein vorausgesagtes Bild (212) zu synthetisieren,  
dadurch gekennzeichnet, daß  
zum Erhalten des vorausgesagten Bilds eine Pixelwert-Interpolation unter Verwendung eines Rundungsverfahrens durchgeführt wird, bei dem durch eine Rundungsverfahrensinformation spezifiziert wird, ob positiv oder negativ zu runden ist, und  
die Rundungsverfahrensinformation aus der Eingangsinformation gewonnen wird.

#### 2. Bilddecoder, der Eingangsinformation dekodiert, mit:

einem Demultiplexer (216), der Bewegungsvektordinformation aus der zu dekodierenden Eingangsinformation gewinnt;  
einem Speicher (209) zum Speichern eines Referenzbilds (210), das ein zuvor dekodiertes Bild darstellt,  
einem Synthesizer (211) zum Synthetisieren eines vorausgesagten Bilds (212), indem eine Bewegungskompensation unter Verwendung der gewonnenen Bewegungsvektordinformation (202) und des Referenzbilds (210) durchgeführt wird,  
dadurch gekennzeichnet, daß  
der Demultiplexer zusätzlich Rundungsverfahrensinformation (1704) aus der Eingangsinformation gewinnt, die spezifiziert ob bei einer Pixelwert-Interpolation in der Bewegungskompensation positiv oder negativ zu runden ist, und  
der Synthesizer Pixelwert-Interpolation unter Verwendung des durch die gewonnene Rundungsverfahrensinformation spezifizierten Rundungsverfahrens durchführt.

3. Computer-läsbares Medium zum Anweisen eines Computers, das Bilddekodierverfahren nach Anspruch 1 auszuführen.

#### 4. Erfindung nach Anspruch 1 oder 2, wobei

die Pixelwert-Interpolation unter Verwendung der positiven Rundung nach folgenden Gleichungen berechnet wird:

$$lb = [(La + Lb + 1)/2];$$

$$lc = [(La + Lc + 1)/2], \text{ und}$$

$$ld = [(La + Lb + Lc + Ld + 2)/4], \text{ und}$$

die Pixelwert-Interpolation unter Verwendung der negativen Rundung nach folgenden Gleichungen berechnet wird:

$$lb = [(La + Lb)/2].$$

$$lc = [(La + Lc)/2], \text{ und}$$

$$ld = [(La + Lb + Lc + Ld + 1)/4].$$

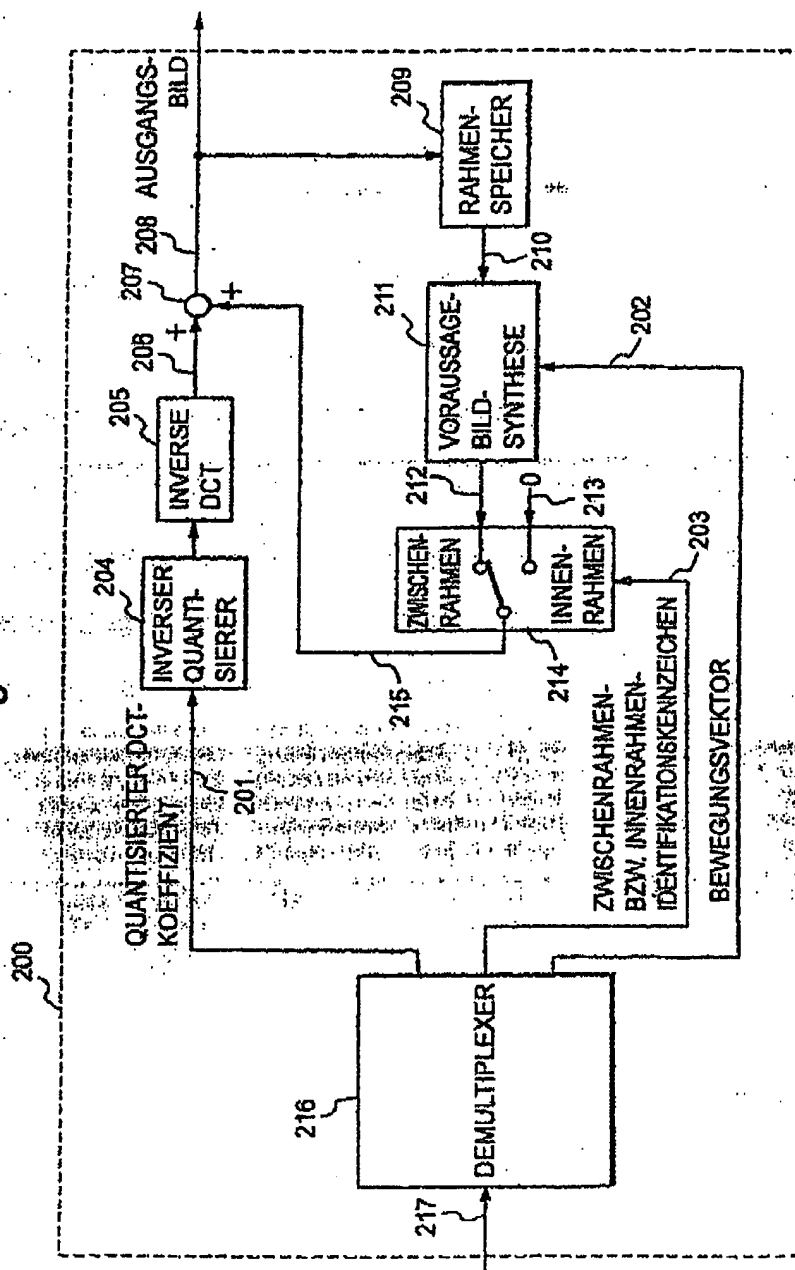
wobei La, Lb, Lc und Ld die jeweiligen Intensitätswerte eines ersten Pixels, eines horizontal an das erste Pixel angrenzenden zweiten Pixels, eines vertikal an das erste Pixel angrenzenden dritten Pixels und eines vertikal an das zweite Pixel und horizontal an das dritte Pixel angrenzenden vierten Pixels sind, und lb, lc und ld die jeweiligen interpolierten Intensitätswerte des mittleren Punkts zwischen dem ersten und dem zweiten Pixel, des mittleren Punkts zwischen dem ersten und dem dritten Pixel und des mittleren Punkts zwischen dem ersten, dem zweiten, dem dritten und dem vierten Pixel sind.

Es folgen 13 Blatt Zeichnungen



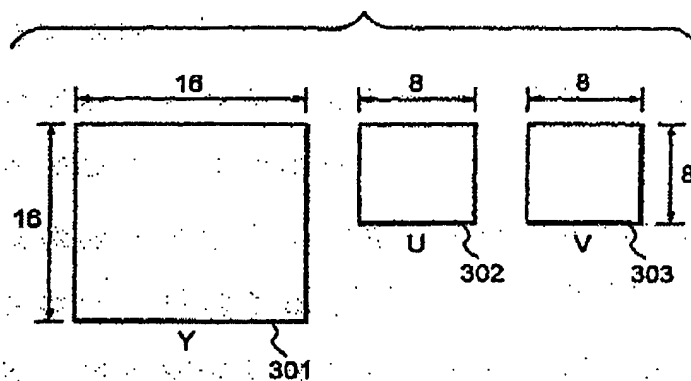
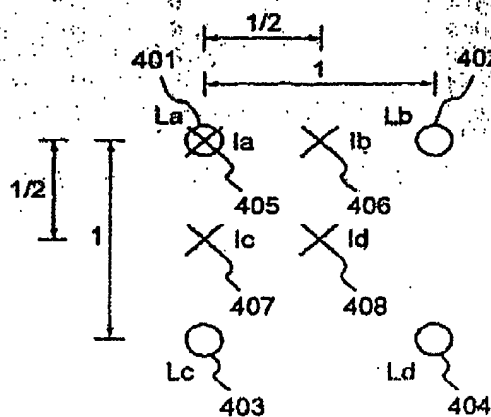
DE 698 17 460 T2 2004.06.09

Figur 2



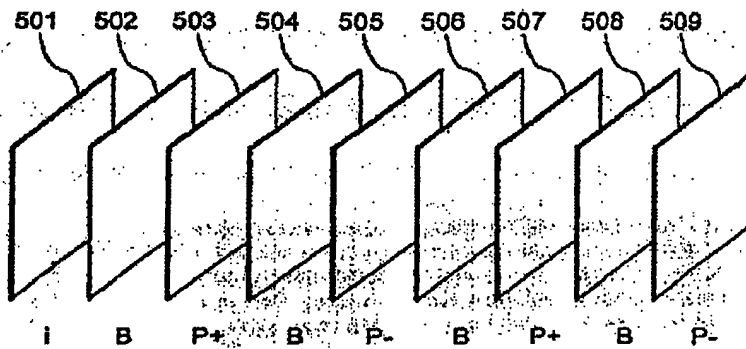


DE 698 17 460 T2 2004.06.09

*Figur 3**Figur 4*

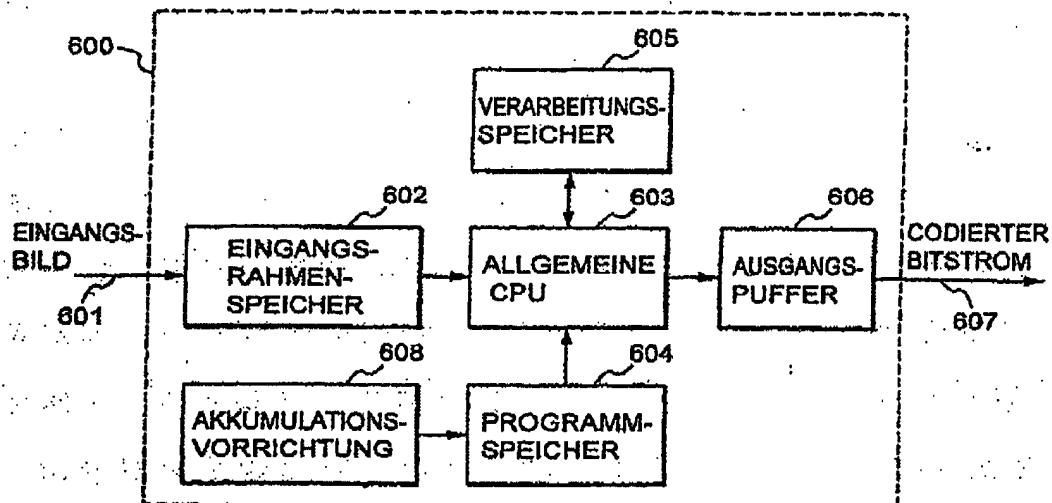
DE 698 17 460 T2 2004.06.09

*Figur 5*

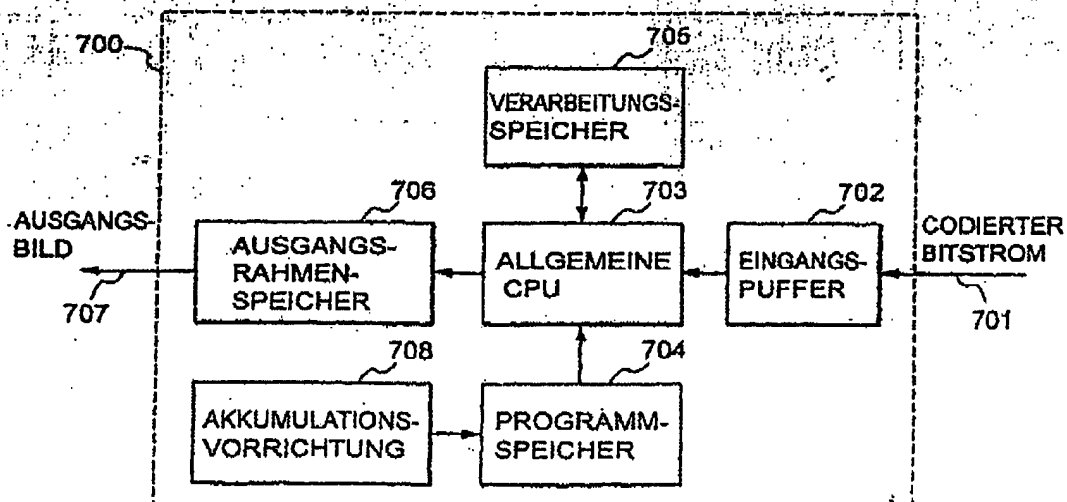


DE 698 17 460 T2 2004.06.09

Figur 6

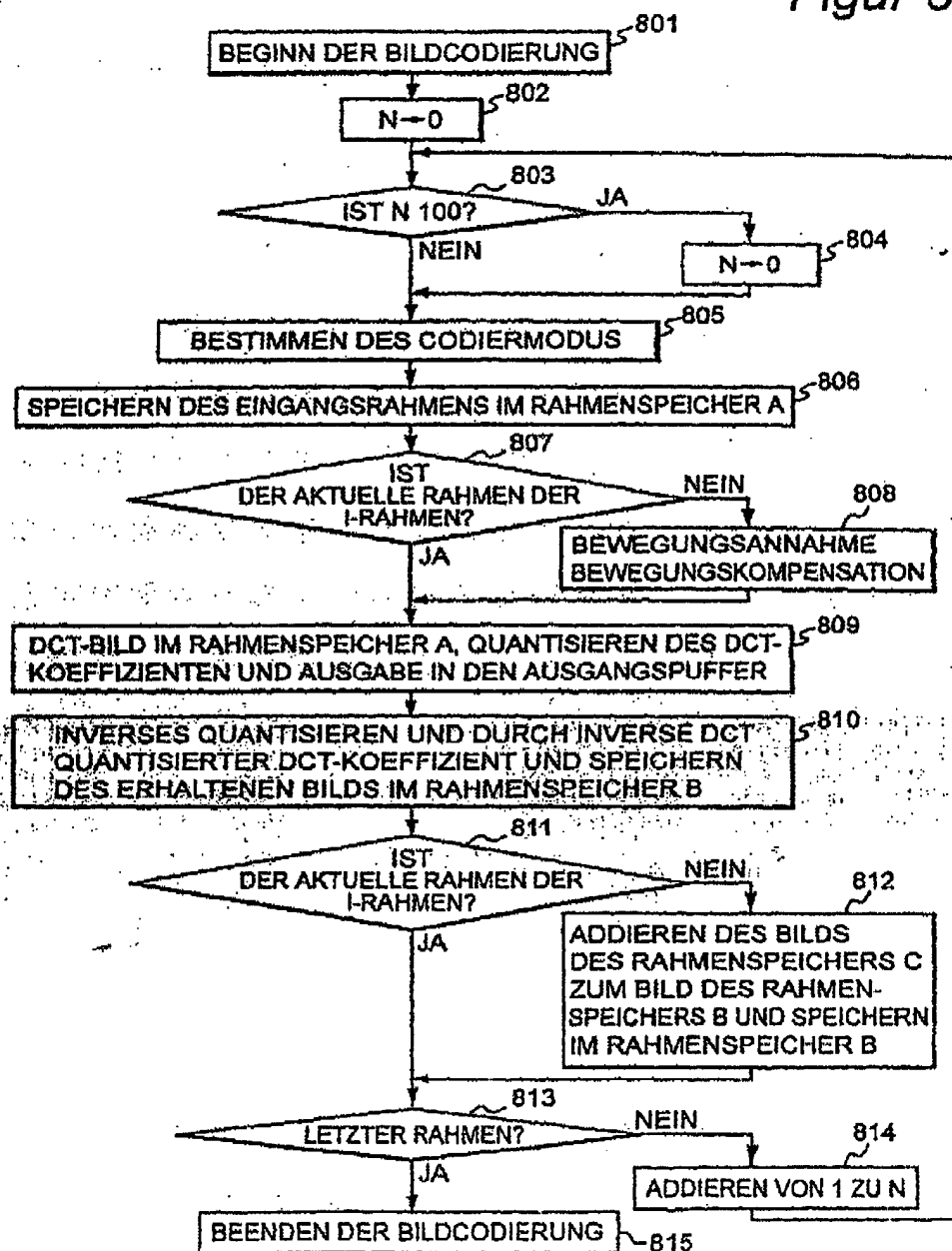


Figur 7



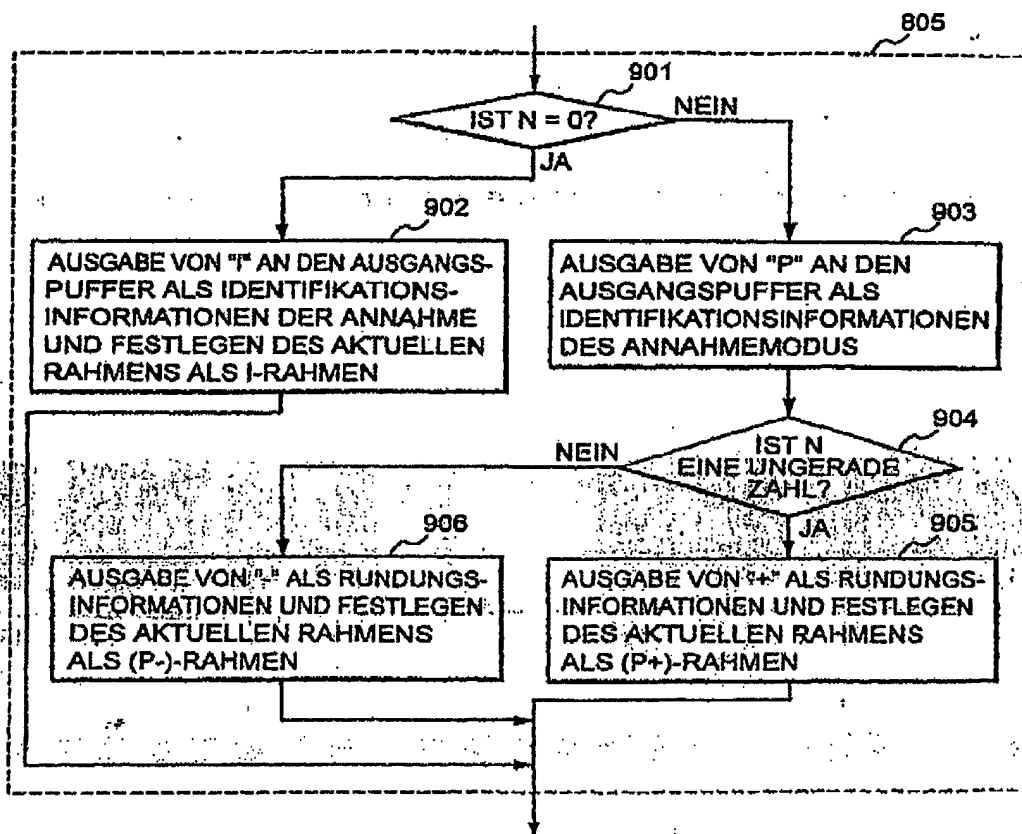
DE 698 17 460 T2 2004.06.09

Figur 8



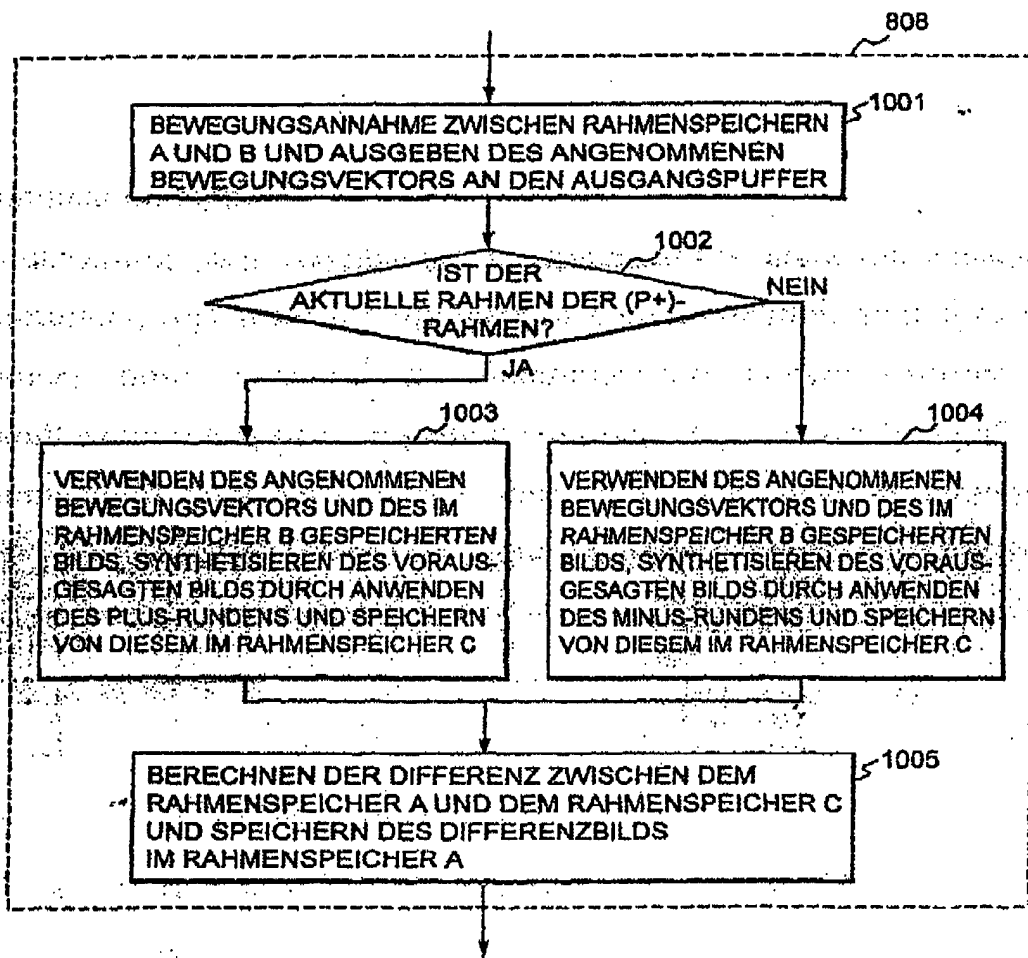
DE 698 17 460 T2 2004.06.09

Figur 9



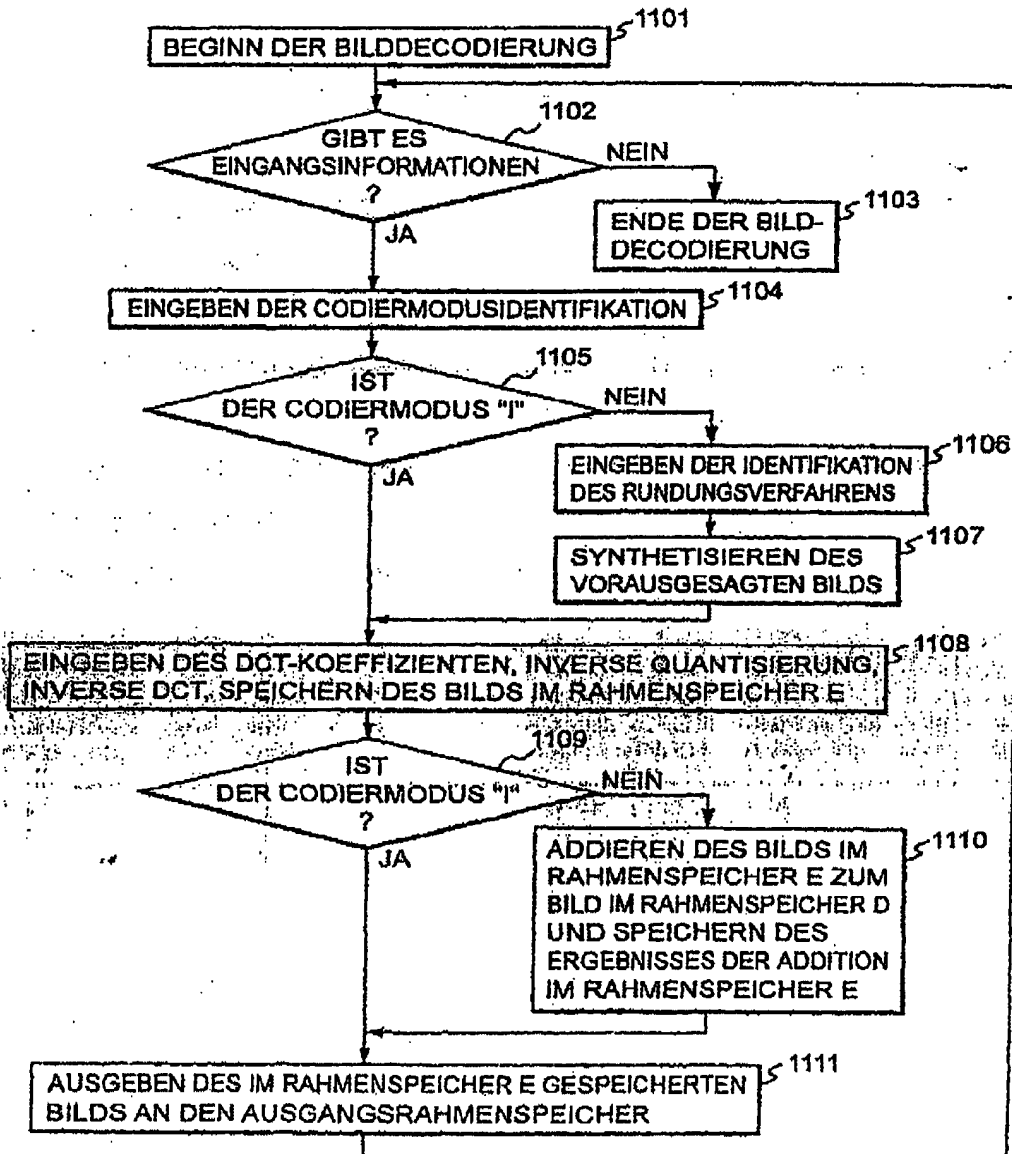
DE 698 17 460 T2 2004.06.09

Figur 10



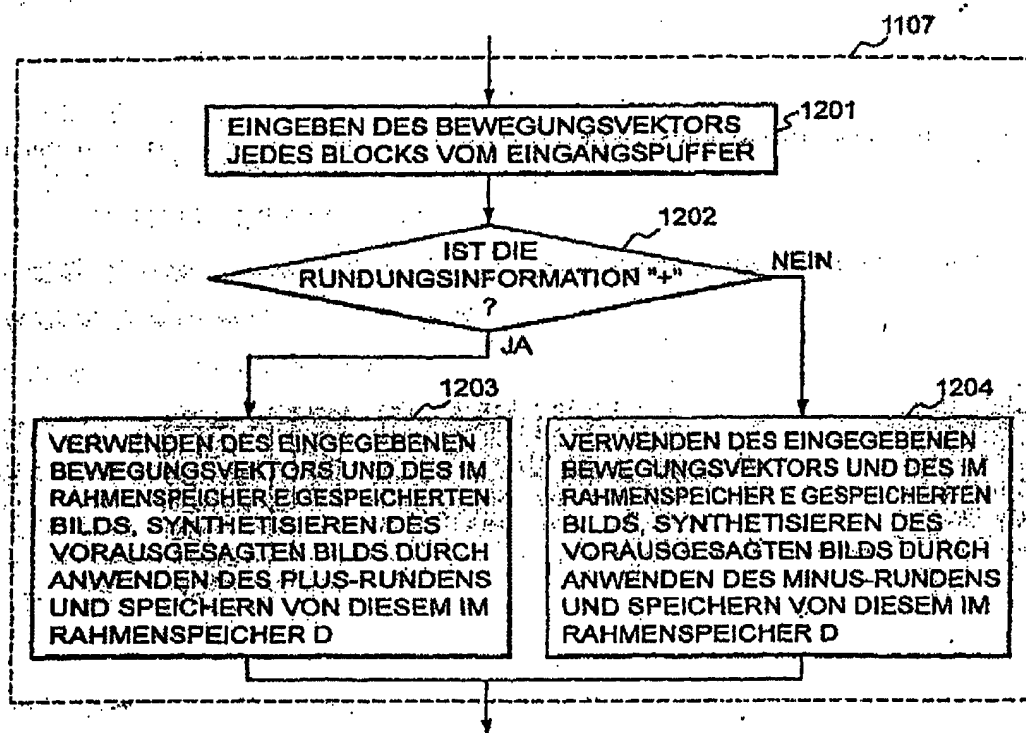
DE 698 17 460 T2 2004.06.09

Figur 11



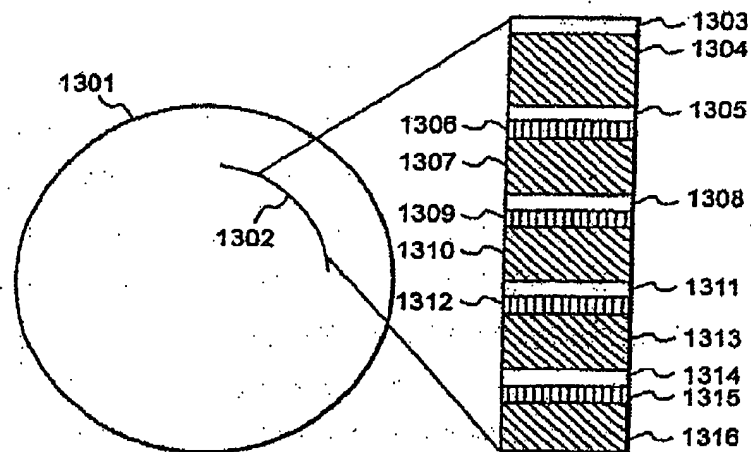
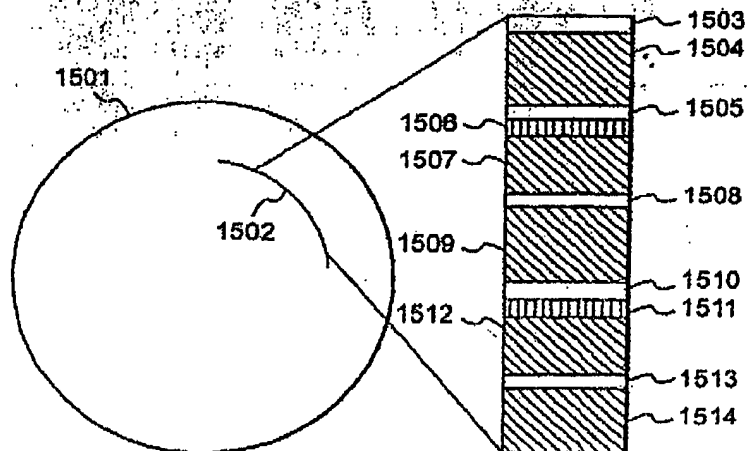
DE 698 17 460 T2 2004.06.09

Figur 12



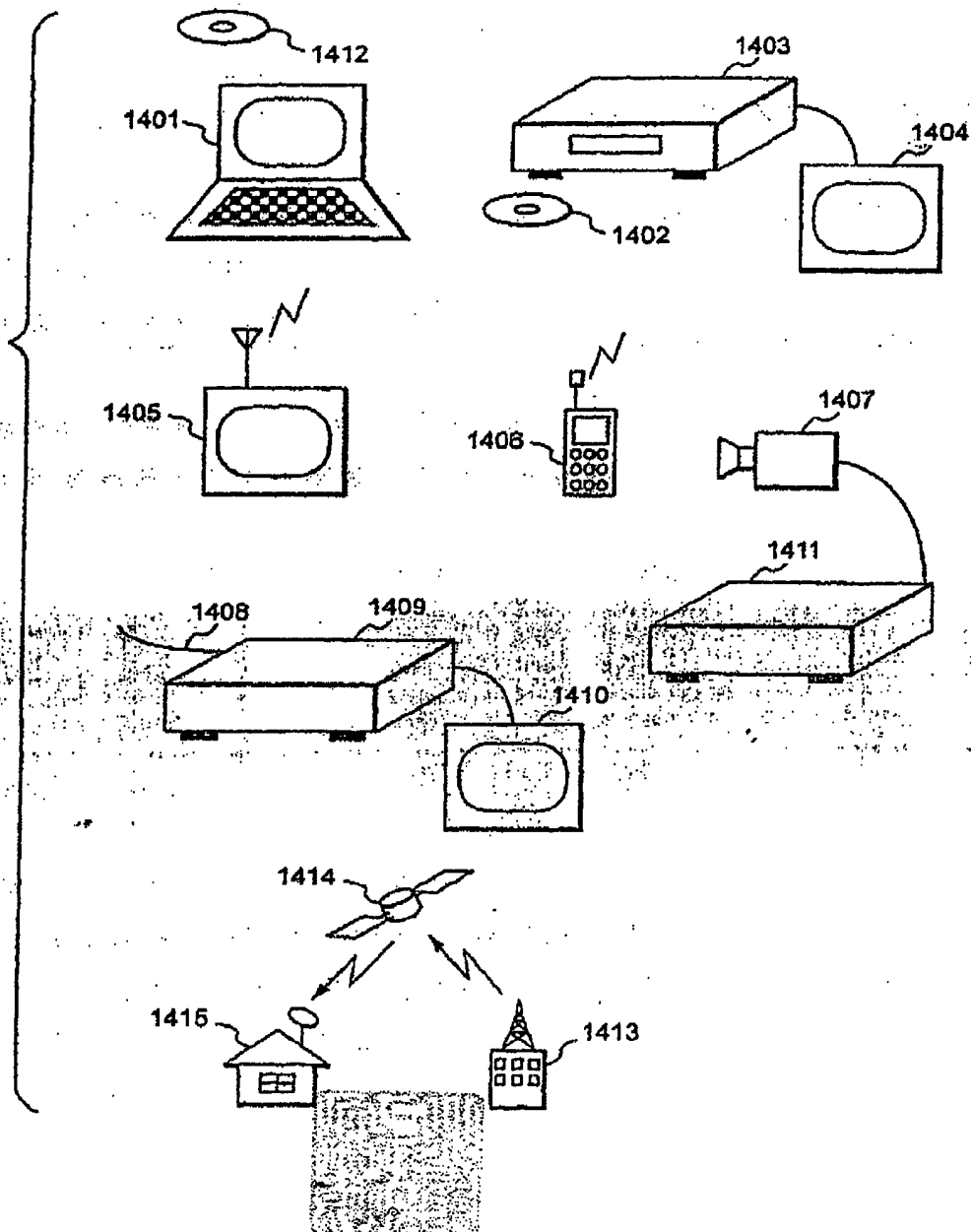


DE 698 17 460 T2 2004.06.09

*Figur 13**Figur 15*

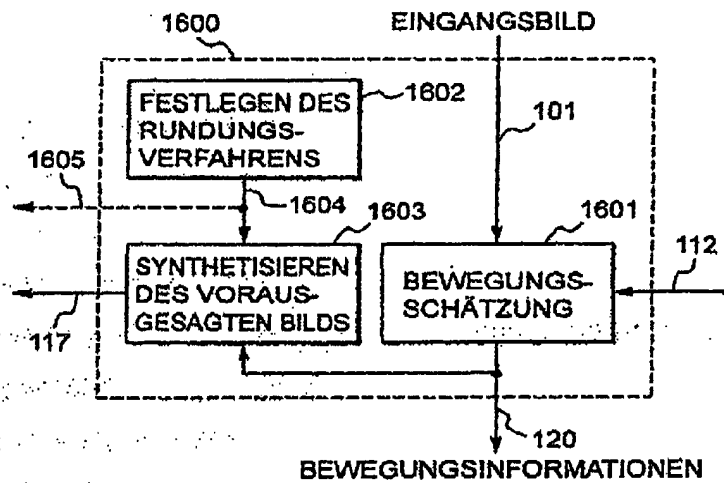
DE 698 17 460 T2 2004.06.09

Figur 14



DE 698 17 460 T2 2004.06.09

Figur 16



Figur 17

